TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Wirtschaftswissenschaften

Professur für Innovationsökonomik

Beyond the Ivory Tower:

Three Essays on the Economics of Publicly Funded Research

Roman Fudickar

Vollständiger Abdruck der von der Fakultät für Wirtschaftswissenschaften der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol.) genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Sebastian Schwenen
Prüfer/in der Dissertation:	1. Prof. Dr. Hanna Hottenrott
	2. Prof. Dr. Michael Kurschilgen

Die Dissertation wurde am 18. April 2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Wirtschaftswissenschaften am 15. Juli 2018 angenommen.

Abstract

This dissertation presents research results on the consequences of public-private knowledge transfers as well as a new method for science evaluations. The first part investigates the role of direct knowledge interactions with public research institutions for the innovation performance of New Technology-Based Firms. The second part explores effects of academic consulting on scientists' research performance. The third part provides a new method for science evaluations which is based on the similarity between scientific texts of individual scientists and externally validated knowledge frontiers. Implications for research and innovation policy are drawn.

Zusammenfassung

Diese Dissertation zeigt Forschungsergebnisse zu den Wirkungen des öffentlich-privaten Wissenstransfers sowie eine neue Methode zur Evaluation wissenschaftlicher Forschungsleistung. Der erste Teil untersucht die Rolle von Wissensinteraktionen mit öffentlicher Forschung für die Innovationsleistung neuer Technologieunternehmen. Der zweite Teil untersucht die Auswirkungen von externer Beratungstätigkeit auf die Forschungsleistung von Wissenschaftlern. Der dritte Teil beschreibt eine neue Methode für wissenschaftliche Evaluationen, die auf der Ähnlichkeit von Textdokumenten zwischen einzelnen Wissenschaftlern und extern validierten Wissensgrenzen beruht. Implikationen für die Forschungs- und Innovationspolitik werden diskutiert.

Acknowledgements

I would like to express my deepest gratitude to my scientific supervisor Hanna Hottenrott who guided and encouraged me throughout all phases of this dissertation and supported me continuously with honest, valuable, and constructive feedback. Thank you for the time you invested in the research project and the great learning I was allowed to experience on this journey.

I also would like to express my gratitude to Piet Hausberg, Cornelia Lawson, Jens Perret, Rohit Prasad, and Michael Zaggl who provided me with great advice for the preparation of this thesis. Moreover, I am very grateful to Michael Kurschilgen for reviewing this dissertation and to Sebastian Schwenen for chairing my dissertation committee.

I also would like to thank my colleagues for the friendly working atmosphere. Particularly, I would like to sincerely thank Irina Baye, Relinde Colen, Thomas Daske, Christoph Gschnaidtner, Florian Inderst, Christian Feilcke, Christian Fischer, Niklas Fourberg, Germain Gaudin, Ali Gibran, Julian Hackinger, Lisa Hamelmann, Mirjam Lange, Anna Lu, Verena Nowak, Florian Peiseler, Tanja Pfaffel, Philipp Pfingstag, Alexander Rasch, Renate Rast, Dennis Rickert, Joel Stiebale, Jan-Philip Schain, Shiva Shekar, Claus Schöttl, Amrei Stammann, Nico Steffen, Magnus Strobel, Manuel Siekmann, Tobias Wenzel, Nicole Wößner, and Tianyu Zhang, for the great experience as a PhD student. Without this supportive as well as fun atmosphere throughout the different challenging phases, this project would have never been successful.

Finally, my deepest and most sincere gratitude belongs to my parents, Gudrun and Wolfgang as well as my sisters Nadine and Marlene. Thank you for your continuous support.

Last, but not least, I would like to thank the institutions that hosted my research, the Düsseldorf Institute for Competition Economics, and the Technical University Munich, School of Management, for providing me with an excellent research infrastructure, a supportive intellectual environment and lots of freedom to pursue my academic interests.

Contents

1. Introduction

> "Governments believe that scientific discovery creates social and economic progress, and so they desire to harness scientific research towards the twin causes of national technological advance and enhanced international competitiveness. In the pursuit of these goals, governments wish to derive maximum utility out of finite public funds while directing the research effort as efficiently as possible. This is the genesis of science governance." (Donovan 2007)

1.1 Motivation

The endogenous growth theory suggests that government investments in new knowledge and human capital are significant contributions to the wealth of nations by means of technical know-how, technology, innovation, and economic growth (Romer 1994; Stephan 1996; May 1997). Such intellectual assets are largely located in universities and public research organizations (Godin and Gingras 2000), and their retention and development is often a key element of national innovation strategies (OECD 2015). Governments invest in their public knowledge infrastructures in order to address market failures associated with the private production of knowledge (Martin and Scott 2000; Edwards 2010; OECD 2016). These investments are usually justified based on the presumption that public knowledge production pays some form of dividends, for instance through knowledge spillovers to the business sector (Jaffe 1998), or more general social returns. In addition to an increasing stock of knowledge, new instruments and methods, knowledge-based firms, and a well-educated workforce (see Salter and Martin 2001 for a review), the economic impact of public research has been attributed to private sector innovation (Dosi et al. 1988; Arora and Gambardella 1994, Aghion and Howitt 2009) and the relevance of public research for innovation and firm performance has repeatedly been stressed (Jaffe 1989; Mansfield 1991; Beise and Stahl 1999).

Mansfield (1991; 1995) was among the first to show for US manufacturing firms that public research contributed to R&D projects such that about 11% of new products and about 9% of new processes would not have been developed without recent academic research. For German manufacturing firms, also Beise and Stahl (1999) conclude that public research has indeed an immediate effect on industrial innovations. Furthermore, the cooperation with external partners, including public research institutions, can lead to a more efficient use of internal R&D resources since "the productivity of own research is affected by the size of the pool or pools it can drew upon" (Griliches 1992). In the study of Cohen et al. (2002) on US manufacturing firms, 31% of firms indicated that public research had made a major contribution to the completion of firms' existing research projects and also on finding new technological opportunities. Moreover, new methods and techniques developed by researchers at public research institutions have been shown to enhance the problem-solving capacities of firms and have the potential to affect R&D outcomes by speeding up experiments, providing more accurate measures, improving results interpretation or re-arranging the research agenda (Dasgupta and David 1994; Antonelli 1999).

Over the past decades, national support for R&D and innovation has witnessed a new rise of public R&D expenditure in many countries, resulting in a worldwide growing research capacity (OECD 2016). Economic indicators in the field of science and technology reflect the level and structure of the efforts undertaken by national governments. Government intramural expenditures on R&D of all OECD countries for example have risen to a total of $136.8 billion in 2015 from $73.2 billion in the year 2000 (all amounts in billions purchasing power parity). The United States and the EU 28 have nearly doubled their public research budget, from $29.1 billion ($25.9 billion) in 2000 to $56.2 (47.5) in the same period. However countries like China and Russia have also begun to fuel their innovation and human capital engine with more research funding to better exploit scientific knowledge as a source for economic growth. Within 15 years, Russian intramural R&D budgets have more than doubled compared to the amount expended in 2000 ($4.7 billion to $11.6 billion) which is now more than the budget of France ($7.9 billion) and the United Kingdom ($3.1 billion) combined (c.f. Spain, $3.7 billion; Italy, $4.0 billion; Germany $16.1 billion; Japan $13.4 billion, all in 2015). The biggest increase, however, can be found in Asia. The Chinese government now spends six times as much on intramural R&D, compared to the year 2000, amounting to $66.1 billion in 2015, which surmounts all other individual countries in absolute terms (presented amounts are based on OECD 2017). R&D expenditure in relative terms, namely as a share of GDP or per capital, obviously provides a slightly different perspective as it takes into account a country's (economic) size. However, this does not obscure the main argument that public research rests high on national policy agendas, and that financial support for public research is perceived as vital investments in future economic benefits.

The OECD (2016) draws several trends that can shape the future of public research systems and their ability to create and disseminate knowledge. First, technology itself is gradually changing the way science is performed, particularly by the growing digitalization, further opening up scientific and technological possibilities. Second, an ongoing expansion of countries' research capacity across the world is likely leading to increased international scientific competition, higher scientific specialization, and a higher demand for knowledge workers. Third, open science movements enable all levels of an inquiring society and especially amateurs to practice scientific research through open data, open access, and citizen science. This indeed also raises expectations among citizens about the contributions of public research to economies and societies. Fourth, the OECD proposes that the main funders of basic research will remain governments in the foreseeable future, but that businesses may increase their financial contribution, reflecting industry's interest in accessing complementary knowledge. A further long-term trend is firms' move away from a "vertical" model of R&D to a "network strategy of innovation" that puts the ability to exploit external knowledge sources, either public or private, at the center of R&D productivity (Foray and Lissoni 2010). As such, publicly-funded knowledge will prospectively gain importance as a complementary source of knowledge to firms' internal R&D and innovation.

The benefits from publicly funded research, and the high national priorities for public basic research underline the importance of understanding the mechanisms of research externalities. Continued economic enquiry of knowledge production and innovation systems is needed for economists, science administrators, and policy makers to maximize the economic returns of investments into publicly-funded research. This dissertation contributes three essays on the knowledge production, dissemination, and economic impact of publicly funded research.

1.2 Literature context

This monography is theoretically grounded in the intersection between literature on the economics of innovation and the economics of science. Subject-matter of studies in these fields are primarily the institutional organization of research, knowledge externalities, and subsequent economic outcomes such as national technological advance and enhanced international competitiveness through innovation.

Scholars of the economics of innovation usually try to comprehend and explain antecedents and consequences of innovation in the public and the private sector. This involves questions about how innovation emerges, how innovation activity drives technological change and economic growth, and how to install and improve policies that foster firms' future innovation success as an intermediate goal for economic well-being and prosperity (e.g. Nelson 1959, Arrow 1962, Mansfield 1968, Kline and Rosenberg 1986, Freeman and Soete 1997).

The literature dedicated to economics of science, in contrast, focuses more on public and private knowledge production systems, scientific productivity, science evaluations, and governing institutions. Typical sets of research questions in this field concern the effective allocation of research funds, reward structures in science, knowledge sharing behavior and secrecy, supply and demand for scientists, scientists' life-cycles, novelty of ideas, knowledge production functions, and the link between science and economic growth (e.g. Merton 1968; Diamond 1996; Stephan 1996, 2012; Partha and David 1994, Lam 2010).

The economics of publicly-funded research are located at the intersection of the economics of innovation and science. While not exclusive to the former, they focus on knowledge externalities from public research to firm R&D and innovation, and the interactions of the two worlds of science and technology. Economic policy typically encourages public research organizations to play an active role in the transfer and commercialization of academic knowledge (Bercovitz and Feldman 2006; Siegel et al. 2007). This, however, has raised questions about the arrangement of scientists' "disinterestedness" in the pursuit of science (Merton 1973) and knowledge transfer activities (Lee 1996). Typical questions that emerge from this conflict of interests are concerned with trade-offs between fundamental research activities and (applied) research commercialization of academic inventions, trade-offs between intellectual property rights and open access / open science, secrecy and withhold of knowledge by scientists, and the overall institutional governance of knowledge transfer activities (e.g. Foray and Lissoni 2010).

1.3 Research aim and summary

The aim of this dissertation is to provide a better understanding of how new knowledge originating in public research institutions is produced, disseminated and how it has an economic impact by addressing three topics concerned with the governance and impact of publicly funded research. In the following, I briefly introduce the three main chapters of this dissertation, present their research objectives, and provide a summary of their findings. A stylized overview of this dissertation is given in Table 1.1.

Chapter 2: The role of public research in the innovation performance of New Technology-Based Firms

The second chapter addresses effects of direct knowledge interactions between public research and New Technology-Based Firms' (NTBFs) innovation performance. NTBFs are generally praised as important agents of technological diffusion due to their ability to successfully commercialize radically innovative products and services (Autio and Yli-Renko 1998; Storey and Tether 1998). These firms have been shown to exhibit higher innovation and growth potentials than non-technology firms (Almus and Nerlinger 1999). At the same time, however, such new firms are constrained in their financial and human resources (Storey and Tether 1998) while trying to compete in fast-changing and R&D-intensive industries that require continuous investments in skilled R&D personnel, and equipment with high asset specificity. This is often referred to their liability of newness and smallness (Stinchcombe 1965; Baum et al. 2000). To circumvent these resource constraints, entrepreneurial firms have been shown to increasingly rely on external knowledge developed by public researchers to complement their R&D processes and to successfully create new markets.

A key question that arises from these observations, and which is addressed here, is "**Are NTBFs with knowledge interactions to public research more likely to introduce new products and services to the market?**". The relationship between knowledge interactions and innovation typically depends on how firms interact with public research and how well they can absorb external knowledge for their R&D (Cohen and Levinthal 2000). Following Grimpe and Hussinger (2013), we distinguish between formal and informal interactions by the presence of a contract for the underlying interaction to separate formal (student internships, supplier relationships, advanced staff training, joint research, and contract research) from informal interactions (personal contacts, scientific conferences). This distinction allows us to investigate the question "**Are there**

differences between formal and informal types of interactions?". Furthermore, firms' absorptive capacity and the persistency with which they interact might moderate this relationship. We address these moderators by asking, "**Does interaction persistency or internal R&D matter for exploiting external knowledge?**".

These questions are addressed by investigating the role of direct interactions with public research institutions for the innovation success of a large sample of NTBFs in Germany. We find that those firms engaging in formal and informal knowledge interactions are more likely to introduce new products and services to the market. The strength of this association, however, depends on the interaction persistency, internal R&D activity, and founders' academic backgrounds. Non-academic start-ups benefit more from continuous informal interactions if they perform their own R&D, which suggests that absorptive capacity matters. In academic start-ups, higher intensities of both formal and informal interactions are associated with a higher innovation likelihood. Moreover, continuous informal interactions complement formal ones in the absence of firms' own R&D activity. We argue that policy makers should encourage knowledge interactions between firms and public research institutions, especially on a continuous basis.

Chapter 3: Academic consulting and individual scientists' research performance

The third chapter aims to assess the positive or negative effects of academic consulting on individual scientists' research performance. Academic consulting is a widespread form of professional advisory service performed by full-time researchers who apply their professional or scholarly expertise outside of their academic institution, often – but not always – for financial compensation (Perkmann and Walsh 2008; Amara et al. 2013). Academic consulting has been shown to be highly valued among industry and government (Cohen et al. 2002; Bekkers and Bodas Freitas 2008; Haucap and Thomas 2014) and deemed an important knowledge transfer channel of public research. While some earlier studies have found positive relationships between academic consulting and research performance (Rebne 1989; Mitchell and Rebne 1995), more recent evidence has warned that consulting activities might come at the cost of reduced research output (Manjarres et al. 2009; Rentocchini et al. 2014). However, explicit evidence for consulting activities is rare despite its relevance. Starting from this controversy, namely whether and how to promote or restrict academic consulting, we ask "**How does consulting affect research outcomes in terms of publications and citations to these publications?**" and "**Does consulting evoke scientists' to cease publishing?**". For these relationships, it might play a role whether consulting is performed for public or private entities, and so therefore we ask, "**Is there a difference between public and private sector consulting?**".

To address these questions, the chapter investigates the effect of academic consulting to public and private sector organizations on scientists' research performance for a sample of social, natural, and engineering science academics in Germany. In contrast to previous research that suggested consulting activities might reduce research output, our analysis provides a more nuanced picture. Public sector consulting comes with lower average citations, particularly for junior researchers. Moreover, engagement in consulting increases the probability to cease publishing research altogether, particularly for private sector consulting. Furthermore, the probability of exit from academic research increases with the intensity of consulting engagement for those at the start or towards the end of their academic career. We draw lessons for research institutions and policies about the promotion of academic consulting.

Chapter 4: Research at the frontier of knowledge: The use of text similarity indicators for measuring scientific excellence

The identification of scientific excellence is of crucial interest to public science administrations that aim to allocate scarce research funds to the most promising projects and persons. Excellent scientists have high probabilities to contribute significantly to science by means of original ideas, findings, and pioneering work, thus pushing back research frontiers and opening up new fields of knowledge (Tijssen et al. 2006). Peer evaluation is, despite its costliness (Rowland 2002) or potential biases (Lee et al. 2013), regarded as the most reliable way to identify excellent scientists (Chubin and Hackett 1990). Such peer evaluations are typically augmented with bibliometric indicators, especially publication and citation counts, but also with content-based analyses (CBA) of publication records. While content-based analysis of scientific publications is on the rise in many fields, especially in biomedicine, economic studies have rarely explored the potential of content-based indicators for science evaluations.

Starting from the continuing need to evaluate scientists and their institutions, we ask "**Does text similarity between individual scientists and externally validated experts provide an alternative indication for scientific excellence compared to citation counts?**". Furthermore, we explore the relationship between text similarity and other research quality indicators. We therefore ask "**How much do research quality indicators like research budget, academic rank or institution rank explain text-based similarity indictors?**".

This study compares citations as a standard measure of scientific excellence to text-based similarity indicators by using natural language processing (NLP) techniques. The proposed text-similarity indicators are based on the idea that scientific proximity between individual scientists and verified knowledge frontiers can be traced through text-based similarity between scientific documents. We test this idea by using co-word analysis to calculate similarity scores for a sample of 1884 international scientists and two knowledge frontier definitions: academic prizes and European Research Council grants. Our comparison suggest a high correlation between content-based similarity scores and citation-based indicators, and that both can be predicted by an individual's academic rank, institutional prestige and research funding. We find that the frontier definition based on academic prizes has a higher explanatory power in our models, compared to the funding frontier. Although a variety of cases can be imagined where text similarity does not reflect a similar scientific alignment (or even quality), we argue that given the “right” reference points, pre-processing and parameters - text similarity approaches can be valuable to complement peer review and standard bibliometric indicators, especially when citation measures may be less meaningful. This is potentially the case for younger scholars since their citation numbers had less time to accumulate.

Table 1.1: Overview of the research findings

	Chapter 2	Chapter 3	Chapter 4
Topic	Academic entrepreneurship	Academic consulting	Science evaluation
Title	The role of public research in the innovation performance of New Technology-Based Firms	Effects of public and private sector consulting on academic research	Research at the frontier of knowledge: The use of text similarity indicators for measuring scientific excellence
Research Questions	• To what extent do public-private knowledge interactions contribute to firms' innovation performance? • How do the rate and intensity of interaction and the presence of internal R&D activities moderate this relationship?	• How does consulting activity of public scientists affect their research performance? • Are there differences between public and private sector consulting?	• Can text-based similarity indicators between publications of individual scientists and documents of validated knowledge frontiers be used to evaluate scientific excellence? • How much do such similarity indicators correlate with bibliometric and non-bibliometric measures?'
Theories	• public-private knowledge interactions • absorptive capacity • knowledge interactions	• detrimental effects of academics' external engagement • research productivity and impact • knowledge transfer channels	• identification of scientific excellence • knowledge frontiers in science • content-based analysis
Methods	• probit regression for 2800 NTBFs in Germany • confirmatory factor analysis • accounting for the self-selection of scientists into knowledge interactions by Heckman correction	• probit and and simultaneous probit regression for 951 academic in Germany • endogenous switching models	• co-word analysis of publications of 1784 international scientists and two knowledge frontier definitions (academic prizes, ERC funding) • correlation analysis and OLS regression
Contributions	• the impact of direct knowledge interactions with public research on innovation success of NTBFs is investigated • formal and informal knowledge interactions are distinguished • the innovation performance of academic and non-academic NTBFs is distinguished • the moderating roles of absorptive capacity and interaction persistency are examined	• new insights on academic consulting (measured by academics' time distributions) and its impact on research dissemination • the effect of consulting activities on (temporary) exit, scientific productivity (publication numbers) and impact (citation numbers) is investigated • distinguish between public and private sector consulting, academic rank, field	• propose a new empirical method to identify scientific excellence based on scientific publication contents • introduce two knowledge frontier definitions based on academic prizes and prestigious grants as a benchmark for scientific excellence • compare 8 different similarity measures • validate the use of content-based indicators with respect to citation-based and independent research quality indicators
Findings	• publicly funded research positively affects NTBFs innovation performance • the majority of sampled NTBFs maintain some contact with PRI, contacts are more common for academic start-ups and firms that pursue internal R&D • both informal and formal interactions increase in NTBFs' probability of radical innovation • magnitude of the effects depend on interaction persistency, internal R&D and founders' academic background • innovation performance can benefit from interaction with PRIs even in absence of own R&D	• reject concerns related to a potential detrimental effect of consulting on research disclosure (publication numbers) • consulting increases the probability of (temporary) exit from academic work, strong effect for private consulting, small effect for high shares of public consulting • less productive academics engage more in consulting • no effect for a decline in overall publication numbers • lower average citations per paper for public consulting, particularly by junior and senior researchers • private sector may provide more research spillovers	• document-document similarity between individual scientists' publications and externally validated knowledge frontiers indeed captures research quality to some extent • average text similarity scores correlate positively with citation counts, citations per article, research budget, academic rank and institution rank • correlations hold for both knowledge frontier definitions, with an overall weaker correlation for the funding frontier • correlations of scores and citations drop with age
Co-authors / Status	• co-author: Hanna Hottenrott • published in The Journal of Technology Transfer	• co-authors: Hanna Hottenrott and Cornelia Lawson • published in Industrial and Corporate Change	• co-author: Hanna Hottenrott

1.4 Dissertation outline

The dissertation is divided into five chapters. After introducing the relevance of public research systems, its embedding in the economic literature, and a brief summary in **Chapter 1**, this thesis presents three essays that each address one of the above outlined sets of research questions. **Chapter 2** provides new insights on the interplay between NTBFs' engagement in formal and informal interactions with public research institutions and their innovation performance. **Chapter 3** addresses the link between academic consulting and research productivity. We draw new insights on the potential effects of private and public sector involvement on academic research on scientific productivity, scientific impact, and the (temporary) discontinuation of research production. **Chapter 4** is concerned with the identification of scientific excellence by using content-based indicators. The study uses co-word analysis and two knowledge frontier definitions to provide a new excellence indication that is based on text similarity to externally validated knowledge frontiers. The validity of the new indicators is tested by comparing them to citation counts and a common set of research quality indicators. **Chapter 5** summarizes the findings and gives an outline on future research avenues.

2. The role of public research in the innovation performance of New Technology-Based Firms

Abstract*

Assessing the role of publicly funded scientific research in entrepreneurial ecosystems is of great interest for science and entrepreneurship policy. Knowledge from academic research flows to the private sector through publications, patents, and researcher mobility as well as through direct interactions between founders and researchers at public research institutions (PRIs). New technology-based firms (NTBFs) are generally praised for high innovativeness despite their resource constraints and liability of newness. This study therefore investigates the role of direct interactions with PRIs for NTBFs' innovation success. In a large sample of NTBFs in Germany, we find that those firms engaging in such knowledge interactions are more likely to introduce new products and services to the market. The strength of this association, however, depends on interaction persistency, internal R&D and the founders' academic backgrounds. Non-academic start-ups benefit more from continuous informal interactions if they perform own R&D, which suggests that absorptive capacity matters. In academic start-ups' higher intensities of both formal and informal interactions are associated with higher innovation likelihood. Moreover, continuous informal interactions complement formal ones in the absence of own R&D activity.

Acknowledgements

The authors thank the ZEW Mannheim for access to the survey data, participants at the 4th Doctoral Workshop in Economics of Innovation, Complexity and Knowledge (University of Turin), the Eurotech alliance workshop on Economics of Innovation and Technological Change (École Polytechnique Fédérale de Lausanne), the European Association for Research in Industrial Economics annual conference 2017 (University of Maastricht) as well as the TUM School of Management Economics Research Seminar for helpful comments. We further acknowledge research funding from the German Research Foundation (DFG), grant number HO 5390/1-1.

* This chapter is based on the peer-reviewed publication "The Role of Public Research in the Innovation Performance of New Technology Based Firms", co-authored by Hanna Hottenrott, accepted for publication in The Journal of Technology Transfer Special Issue with the title "Economic, Technological and Societal Impacts of Entrepreneurial Ecosystems".

2.1 Introduction

Knowledge production is largely located in universities and public research organizations (Godin and Gingras 2000). Supporting public research is therefore generally a key element of national innovation strategies (OECD 2015). Governments invest in public research infrastructure in order to address market failure associated with the private production of knowledge (Martin and Scott 2000; OECD 2016) and to foster growth-oriented entrepreneurship (Mason and Brown 2014). Indeed, National System of Entrepreneurship research attributes an important role to scientific research institutions for the production and (free) dissemination of new knowledge that benefits the creation and performance of entrepreneurial ventures (Acs et al. 2014; Autio et al. 2014).

New Technology Based Firms (NTBFs)[1] may particularly benefit from research-based knowledge spillovers. NTBFs typically compete in knowledge-intensive industries marked by rapid technological change that requires continuous investments in skilled personnel and equipment with high asset specificity. Despite their financial and human resource constraints that limit the extent to which they can scale up investment in own research and development (R&D), NTBFs have been shown to have a particularly high innovation and growth potential (Storey and Tether 1998; Almus and Nerlinger 1999). It seems, however, like a paradox that NTBFs are highly innovative despite their resource constraints. Addressing this puzzle, Audretsch (1995, p. 179) argued that knowledge spillovers from the entrepreneurial ecosystem help these firms to innovate:

> 'How are these small and frequently new firms able to generate innovative output when undertaking a generally negligible amount of investment into knowledge-generating inputs, such as R&D? One answer is apparently through exploiting knowledge created by expenditures on research in universities and on R&D in large corporations'.

Subsequent research focused more on the role of the entrepreneurial ecosystem and studied entrepreneurial activity in the local context taking into account a more diverse set of actors and factors that enable productive entrepreneurship (Audretsch et al. 2011). Previous literature therefore stressed the role of public research institutes (PRIs) in such systems of agents (Audretsch and Lehmann 2005; Audretsch et al. 2005; Mueller 2006; Audretsch et al. 2011; Guerrero et al. 2015). For instance, Audretsch and Lehmann (2005) find that the knowledge output of universities positively influences the number of firms located close by and Audretsch et al. (2005) show that NTBFs have a high propensity to locate close to universities, presumably in order to appropriate knowledge spillovers. In line with the Knowledge Spillover Theory of

[1] NTBFs are defined as firms which are i) independently owned, ii) less than 25 years old, and iii) operating in a high-tech or knowledge intensive industry (Ganotakis 2012).

Entrepreneurship, both studies conclude that knowledge spilling over from academic research is an important source of opportunities for entrepreneurship.

Previous research further suggests that knowledge spillovers not only happen prior to the start of a new firm and that they are happening intentionally. New firms actively seek knowledge interactions with PRIs to complement their internal knowledge production with external knowledge from scientific research (Baum et al. 2000; Lynskey 2009; Bellucci and Pennacchio 2016).

Despite its relevance for innovation and entrepreneurship policy, evidence on the effects of such direct knowledge interactions on new firms' innovation performance is still rare. This study therefore explores the extent to which NTBFs make use of public research through different forms of interactions with PRIs and whether these interactions contribute to higher innovation success. In particular, we analyse the founders' engagement in formal and informal modes of knowledge interactions in a sample of more than 2800 NTBFs in Germany. We further investigate whether the extent to which NTBFs benefit depends on the persistency with which they maintain these knowledge interactions and whether the internal knowledge base, developed through their own R&D activities or through the founders' academic background, matters.

Our findings underline the importance of public research in entrepreneurial ecosystems by showing that the majority of sampled NTBFs maintain some form of contact with PRIs. Contacts to PRIs are more common for academic start-ups and firms that pursue internal R&D, but also a considerable share of non-R&D-active start-ups and firms without academic founders interact with PRIs. Innovation outcome models that account for the selection of firms into these knowledge interactions show that both informal and formal interactions are associated with an increase in NTBFs' probability of initiating radical innovation. The magnitude of the effects, however, depend on interaction persistency, internal R&D and the founders' academic background. Estimation of the joint effects of formal interaction, informal interaction and internal R&D confirms that internal R&D is not only an important innovation driver, but also that the benefits of engaging in knowledge interactions with PRIs can be higher for R&D-active firms. However, internal R&D is more important in non-academic start-ups indicating that academic background may substitute internal R&D at least in case of informal interactions. For academic start-ups without internal R&D, continuous informal and formal interaction complement each other.

This study contributes to prior research on the role of public research in entrepreneurial ecosystems by showing that direct knowledge interactions between NTBFs and public research can explain variation in innovativeness. Second, our results suggest that both formal and informal interactions matter, but that the founders' academic background increases the returns to formal interactions. Third, the results confirm the insight from previous research that firms utilize external knowledge better if they have a higher absorptive capacity created through internal R&D (Cohen and Levinthal 1990; Acs et al. 2014). The results, however, also show that for academic start-ups innovation performance can benefit from interaction with PRIs even in absence of own R&D.

The results support the argument that public research plays a key role for knowledge and technology transfer and eventually innovation in entrepreneurial ecosystems. They extend the insights that publicly funded research positively affects firms' innovation performance (Mansfield 1991; 1995; Beise and Stahl 1999; Cohen et al. 2002; Bellucci and Pennacchio 2016) to the context of NTBFs. Furthermore, the results underline that the importance of informal knowledge interactions (Meyer-Krahmer and Schmoch 1998; Cohen et al. 2002) applies also to NTBFs' innovation performance.

The conclusions from this study are relevant for policy makers as they stress the economic impact of universities and public research organizations. Their role as relevant source of knowledge spillovers and as valuable collaboration partners in the entrepreneurial ecosystem draws attention to public research funding as policy tool for fostering technology-based entrepreneurship. The results may also encourage policy makers to promote formal as well as continuous informal knowledge interactions between NTBFs and public research institutions as part of their entrepreneurship policy.

2.2 Publicly funded research and industrial innovation

Publicly funded scientific research is a major contributor to the stock of useful knowledge which, as a public good, benefits many actors in the society and economy including entrepreneurs (Callon 1994; Cohen et al. 2002; Guerrero et al. 2015). Advances in scientific understanding and techniques originating from publicly funded research have even been regarded as "the most powerful and, over the long run, almost certainly the most important source of new technological opportunities" (Klevorick et al. 1995).

Mansfield (1991; 1995) was among the first to show for US manufacturing firms that public research contributed to R&D projects such that about 11% of new products and about 9% of new processes would not have been developed without recent academic research. For German manufacturing firms, Beise and Stahl (1999) conclude that public research has indeed an immediate effect on industrial innovations. Further, the cooperation with external partners, including public research institutions, can lead to a more efficient use of internal R&D resources since "the productivity of own research is affected by the size of the pool or pools it can drew upon" (Griliches 1992). In the study of Cohen et al. (2002) on US manufacturing firms, 31% of firms indicated that public research had made a major contribution to the completion of firms' existing research projects and also in finding new technological opportunities. Moreover, new methods and techniques developed by researchers at PRIs have been shown to enhance the problem-solving capacities of firms and have the potential to affect R&D outcomes by speeding up experiments, providing more accurate measures, improving results interpretation or re-arranging the research agenda (Dasgupta and David 1994; Antonelli 1999).

In addition to supporting innovation in established companies, government-funded research may benefit young entrepreneurial firms (Audresch and Lehmann 20005; Mueller 2006; Audretsch et al. 2011). Entrepreneurship may therefore constitute a channel for utilizing publicly funded research through NTBFs' role as sources of radical innovation and agents of technology diffusion (Autio and Yli-Renko 1998). By providing access to specialized expert knowhow, knowledge from public research may guide the search for innovation by limiting the technological landscape of possible solutions to the most promising areas while eliminating others (Nelson 1982; Fleming and Sorenson 2004, Roper et al. 2017). Young firms are usually constrained in their ability to finance capital intensive R&D (Storey and Tether 1998) and are confronted with the liability of newness and smallness (Stinchcombe 1965; Baum et al. 2000). Such constraints stem from difficulties to compete with large, established companies in the attraction of skilled personnel, lack of credibility, and limited financial resources. The use of external knowledge, and particularly scientific knowledge produced in PRIs, may be a way for NTBFs to cope with these constraints and hence explain the puzzle of their innovativeness. Public research can also influence the quality and pace of innovation by reducing the time lag between the origin of new knowledge and its utilization in novel products and services. By using state-of-the-art results, methods, and equipment residing in universities and PRIs, firms are more likely to produce unprecedented knowledge and eventually inventions that lead to more radical innovations (Tether 2002; Amara and Landry 2005).

NTBFs may have certain advantages over non-TBFs in tapping external sources for knowledge in knowledge-intensive areas such as the high-tech sectors. Founders of NTBFs are likely to have a university degree or practical research experience facilitating them to use their cumulated knowledge and experience to exploit technological opportunities with their firm (Murray 2004; Ganotakis 2012; Guerrero et al. 2015).

Despite these insights, studies on the link between publicly funded research and NTBFs' innovation performance are rare. Baum et al. (2000) study the connection between the composition of start-ups' alliance networks at founding and innovation of Canadian biotechnology start-ups. They suggest that firms exhibit stronger initial (innovation) performance if they form and configure alliances with external partners, including PRIs. Ganotakis and Love (2012) investigate factors that lead to the use of different knowledge sources, the relationships between those sources of knowledge, and the effect that each knowledge source has on the innovative activity of NTBFs. For a sample of UK NTBFs, they find that while customer/supplier collaboration has a positive impact on innovation decision and success, collaboration with universities and research institutions does not seem to drive their innovation performance. This is in contrast to a study on technology-based start-up firms in Japan by Lynskey (2009) who finds a positive link between knowledge spillovers from PRIs through patents or scientific publications and innovation to NTBFs.

Public-private knowledge spillovers

The knowledge produced by PRIs can be explicit/codified in the form of publications in peer-reviewed journals, patents or licenses, or implicit/tacit including know-how, techniques or methods. While it is relatively straightforward for knowledge users to access explicit knowledge through markets for knowledge (Antonelli 1999, Grimpe and Sofka 2016), implicit knowledge typically requires some form of interaction between the provider and the user. Implicit or tacit knowledge, like inherited practices, implied values and prejudgments, is difficult to transfer between individuals (and organizations) by means of formalization or verbalization, and often requires direct cooperation in the form of what can be labelled direct knowledge interactions. Such direct knowledge interactions have higher levels of natural excludability which refers to those parts of knowledge that is transferred best through working with discoverers (Zucker et al. 1998). Direct knowledge interactions may therefore provide a greater knowledge-based competitive advantage compared to codified information channels (Grimpe and Hussinger 2013).

Among the direct knowledge interactions, we can distinguish formal and informal types. Despite the fact that informal interactions are difficult to quantify, they have been acknowledged as being highly relevant and as a prerequisite for subsequent formal knowledge exchanges. Cohen et al. (2002) report that informal interactions and conference meetings are rated second in importance out of 10 possible sources[2] of public research knowledge across several US industries. Similarly, Siegel et al. (2004) highlight informal transfer of know-how to be the second most frequently mentioned output by managers and entrepreneurs. Ponomariov and Boardman (2008) further show that university scientists' involvement in informal interactions with private sector companies increases the probability of undertaking (formal) collaborative research with industry, and also the time allocated to collaborative research with industry. Olmos-Peñuela et al. (2014) qualitatively analyze collaborations between universities and companies and conclude that formal and informal collaborations can co-exist and even strengthen one another. Their interviews with Spanish scientists also highlight that informal interactions often lead to formal agreements. They further study scientists' motivational differences for engaging in formal or informal interactions and find that formal agreements are made when a collaboration requires significant economic resources and when there is a need to formalize the conditions under which a specific project is

[2] Other public knowledge sources were (in percent): publications and reports (41.2), informal interactions (35.6), meetings or conferences (35.1), consulting (31.8), contract research (20.9), recent hires (19.6), cooperation and joint ventures (17.9), patents (17.5), licences (9.5), and personnel exchange (5.8).

carried out. Grimpe and Hussinger (2013) distinguish formal and informal by the presence of a contract for the underlying interaction. Their findings suggest that the innovation performance of manufacturing firms in Germany is highest when firms engage in both, formal and informal interactions and therefore conclude that there is a complementary relationship between these direct interaction types.

Absorptive capacity and interaction persistency

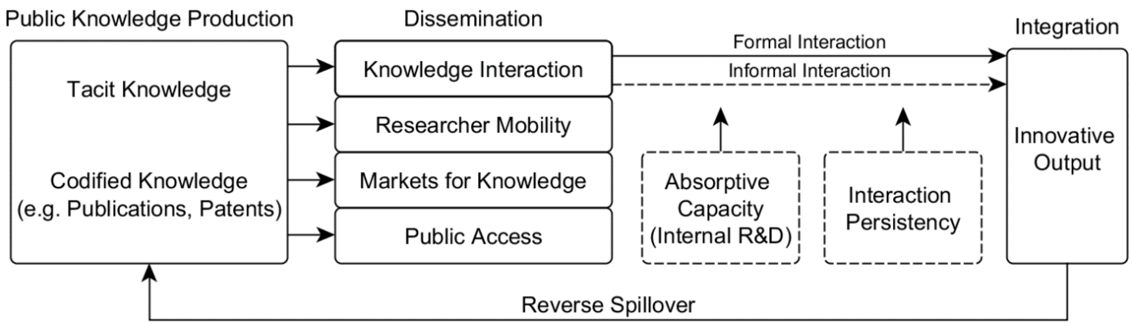
The ability to evaluate, acquire and utilize external knowledge, however, depends not only on the type of knowledge interaction. It also depends on the firm's ability to integrate external knowledge into new products, services or processes (Gambardella 1992). Absorptive capacity as such is not directly observable, and rather "largely a function of the firm's level of prior related knowledge" which enables them to better understand, replicate and build on external knowledge (Cohen and Levinthal 1990). Building on this theory, Qian and Acs (2013) introduce the notion of Entrepreneurial Absorptive Capacity, that focuses on "the ability of an entrepreneur to understand new knowledge, recognize its value, and subsequently commercialize it by creating a firm". The authors highlight two types of knowledge that need to be absorbed by new ventures to successfully create and operate a new firm, i.e. scientific knowledge to understand and to determine the market value of new combinations of knowledge, and market or business knowledge to create value from inventions. An academic background, i.e. a university degree, may also constitute a determinant of absorptive capacity by enabling the academic entrepreneur to better utilize knowledge spillovers and recognize opportunities for learning and collaboration. The moderating role of absorptive capacity in the link between external knowledge and innovation performance has been addressed in a few studies. These studies frequently, but not always, confirm a certain complementarity between internal and external knowledge (e.g. Mansfield 1995; Beise and Stahl 1999; Vega-Jurado et al. 2009; Brehm and Lundin 2012; Higón 2016). Brehm and Lundin (2012), for example, show for Chinese manufacturing firms that universities' impact on commercial innovation depends on the types of activities they perform, especially on their own R&D efforts. In contrast to above findings, Vega-Jurado et al. (2009) find that for Spanish manufacturing firms, cooperation with scientific agents does not constitute a key factor in developing new products, especially when firms put a lot of effort into developing in-house R&D activities.

Thus, the ability to absorb external knowledge may further depend on the collaboration continuity or persistency of interaction. Repeated and even continuous interaction between collaborators can foster the exchange of tacit and codified knowledge, increase the efficiency of collaboration (Nieto and Santamaria 2007) and build-up trust among collaborators (Gulati 1995). Belderbos et al. (2015) for example address specific temporal patterns of R&D collaboration and find that persistent collaboration with all types of partners is associated with higher innovation performance. In the context of radical innovation, it has been argued that sustained and intense interactions between firms and external sources of technical information increase the likelihood of this information being used (Amara and Landry 2005; Nieto and Santamaría 2007). Similarly, Godoe (2000) concludes that "intimate and prolonged interaction" leads to higher quality innovations.

Hypotheses

From the arguments in the previous sections, we derive the theoretical framework presented in Figure 2.1.[3] The framework addresses the research question: to what extent do public-private knowledge interactions contribute to firms' innovation performance considering the rate and intensity of these interactions and the presence of internal R&D activities?

Figure 2.1: Knowledge flows from public research to industrial innovation



The first hypothesis addresses the direct link between knowledge interactions and NTBF innovation as discussed in subsection 2.1. First, interactions with PRIs may provide specific answers or solutions to firms' R&D-related challenges and thereby increase innovation output. Second, such knowledge interactions may expand firms' research-related skills and methods affecting the way firms conduct research, which may ultimately affect the translation of R&D into innovation. A higher interaction intensity, in the sense of using more of different types of interactions, should likewise contribute to innovation performance. We therefore expect that

Hypothesis 1: Knowledge interactions with PRIs are positively associated with NTBFs' innovation performance.

In addition, the persistency of the interaction with external partners may positively influence firms' innovation outcomes (Belderbos et al. 2015). As argued in subsection 2.2, the more persistently a firm interacts with PRIs, the more tacit knowledge can potentially be transferred between the parties. For example, a firm's researcher who repeatedly visits scientific conferences

[3] It should be noted that the theoretical framework also includes other knowledge transfer channels besides direct interactions as well as potential reverse spillovers from industrial innovation to public knowledge production (Perkmann and Walsh 2009; Hottenrott and Lawson 2014; Sohn 2014) which are, however, beyond the scope of the following empirical analysis.

related to his problem set has a much higher chance of accumulating useful knowledge and ideas from speakers and peers than someone who chooses to visit conferences only occasionally.

Another argument relates to the improved quality of interaction due to temporal continuity. The more often two parties interact, the better they get to know each other, the less frictional losses they incur because communication is more effective, and the more experience is gained at managing collaborations and transferring knowledge (Powell et al. 1999). This fine-tuning between collaborators over time increases the collaboration efficiency and results in more knowledge being transferred and acquired. For example, after several joint research projects completed together, both parties, firm and public scientist, know much better how to handle each other’s expectations, timeliness, or perception of precision, than is the case for occasional interaction. From these two persistency-related arguments, we derive the following hypothesis:

Hypothesis 2: The association between knowledge interactions with PRIs and innovation performance is stronger for continuous interactions than for occasional ones.

As also illustrated in Figure 2.1, the extent to which knowledge interactions affect innovation outcomes may be determined by firms’ absorptive capacity. An often-used measure for this capacity is internal R&D (Lucena and Roper 2016). A firm that builds up internal R&D facilities and employs its own R&D personnel has, according to the theory, a better chance of leveraging external knowledge while firms without internal R&D lack an understanding of the value of knowledge and ways to integrate it into their own innovation processes. We therefore expect that firms that invest in R&D benefit more from knowledge interactions with PRIs than firms without internal R&D:

Hypothesis 3a: The association between knowledge interactions with PRIs and innovation is stronger in R&D-active firms.

For academic entrepreneurs, however, the academic background may substitute for current internal R&D. Academic start-ups may exploit the insights, research experience, and even results produced in academia (Lockett et al. 2005). For academic start-ups, we therefore expect that the importance of internal R&D as a catalyst for knowledge spillovers from public research is smaller than in other firms. More precisely, we hypothesize that

Hypothesis 3b: The association between knowledge interactions with PRIs and innovation depends less on R&D in academic start-ups than in non-academic ones.

As argued in subsection 2.1, a firm’s innovation performance might also depend on the type of interaction, i.e. formal or informal. Formal interactions involve clearly defined goals (or tasks) and some form of financial compensation or commitments on the part of the firms. For firms, formal engagement with public research is an investment (with opportunity costs) that is monitored and even legally enforced if outcomes are inadequate. Scientists engaging in formal interactions have an obligation to deliver a certain desired outcome for the firm, in order to retain and increase reputation, and, in the worst case, to avoid potential legal claims. Legal instruments like contracts or cooperation agreements therefore require hard commitments with risks and opportunities for both parties.

Informal interactions in contrast can be seen as softer commitments since there is usually no legally binding agreement to cooperate until a desired outcome is achieved. These loose commitments might also be goal-oriented but they rely on the goodwill of the scientist and their willingness to share knowledge rather than on any (legal) obligation. Informal interactions can therefore be regarded as more inspirational, open-ended and less professionally managed. Firms potentially gain specialized knowledge from informal interactions with public research while scientists on the other hand enjoy greater autonomy in informal interactions, since they can apply their knowledge, and gain exposure to the latest trends, applied problems and unsolved research questions but are free to abandon or postpone the cooperation if they wish. Based on this distinction between formal and informal types of interactions, we expect that firms investing in internal R&D to pursue clearly defined goals benefit more from formal interactions than from informal interactions:

Hypothesis 4: The association between formal interactions and innovation performance depends more on internal R&D than the association between informal interactions and innovation performance.

While in principle this hypothesis stands for both academic and non-academic start-ups, one may again argue that for the former, internal R&D can be substituted by the founder's academic background and current internal R&D should matter less for the returns from both formal and informal interactions.

The theoretical framework also incorporates potential complementarity between informal and formal interactions like identified in previous studies (Grimpe and Hussinger 2013). In general, we would expect that firms using more diverse interaction types increase their chances of extracting new knowledge helping their innovation search (Catozzella and Vivarelli 2014). Hence, a firm that engages in both formal and informal interactions can potentially absorb more diverse knowledge and exploit synergies between different inputs compared to a firm that uses only one approach. The benefits from using both formal and informal interaction may, however, further depend on the firm's internal R&D. In particular, if firms conduct internal R&D, and therefore are less in need of translational interaction, the value of additional informal knowledge exchange may be smaller if the firm also engages in formal interactions. This may be particularly true for academic founders who are trained to understand scientific communication and methods. Informal and formal modes of interaction may therefore be substitutes rather than complements in terms of innovation outcomes. On the contrary, for non-R&D performers, the value of informal knowledge interaction may increase when formal interaction intensity is higher. Thus, for non-R&D performers, the usefulness of informal interactions may depend on the intensity of formal interactions. In other words, informal modes of interaction complement formal ones. Thus, we hypothesize that:

Hypothesis 5: The marginal benefit of informal interactions is larger at higher intensities of formal interaction in non-R&D performing NTBFs and is smaller for R&D-performing ones.

2.3 Data and model specification

Data source

The following analysis is based on firm-level data of German technology-based firms established between 2001 and 2006. The data had been collected as part of the ZEW High-tech Survey using computer-aided telephone interviews (CATI). The participating firms were drawn as a random sample from the KfW/ZEW start-up panel stratified by sector and founding year (Fryges et al. 2010). Each of the six annual start-up cohorts (2001-2006) comprises around 14-19% of the sample. Around 33% of the firms belong to manufacturing sectors such as chemicals, pharmaceuticals, engineering, electronics or telecommunication and the remaining firms belong to technology-based services (33%) or knowledge-intensive services (34%) sectors. These firms qualify as NTBFs which are defined as i) independently owned, ii) less than 25 years old, and iii) operating in a high-tech industry (either products or services) (Ganotakis 2012). We can further distinguish academic entrepreneurs in the broader sense from other start-ups by considering start-ups in which at least one of the founders holds a degree from a university or technical college or is currently enrolled as "academic". A list of variables and pair-wise correlation matrices are provided in Tables 2.7 and 2.8 in the Appendix. After excluding incomplete responses, the final sample comprises information on 2879 NTBFs of which 68% are classified as academic start-ups.

Measurement

Innovation success

The outcome variable of interest is a binary indicator of whether a firm had introduced a product innovation that was entirely new to the market (*market novelty*). Since all firms in the sample are new, this question refers to the time elapsed since the company was founded. In the sample more than one third of the firms (37%) report to have introduced such a market novelty confirming the high potential for radical innovation typically attributed to NTBFs. Academic start-ups are more likely to have introduced a market novelty (42%) compared to start-ups without academic founders (28%, see Table 2.2 for the mean differences between academic and non-academic start-ups).

Knowledge interactions and internal R&D

To measure the nature and the extent of knowledge interactions with PRIs, the survey respondents were asked: "In the context of your operating activities, does your company have any form of contacts with universities, or other scientific institutions?". If the respondents answered yes, they were asked to further specify those contacts from a given (order-randomized) list and in addition to disclose the persistency of this interaction type, i.e. occasional or continuous. The sub-questions were 1) "Did you conduct joint research projects together with scientific institutions?", 2) "Do you award contracts for research or consulting to public research institutions?", 3) "Are public research institutions your customers?", 4) "Do your employees receive advanced training at public research institutions?", 5) "Do you employ students for internships and final theses in your company?", 6) "Do you maintain personal informal contacts to public research institutions?", 7) "Do you visit scientific conferences or congresses?".

Correlation analysis shows that the knowledge interaction variables are highly correlated (see Table 2.8 in the Appendix). To address this issue, we use maximum likelihood factor analysis to confirm the presence of two main latent factors: *formal interactions* including joint research, supplier relationships, contract research, advanced staff training, and student internships; and *informal interactions* including personal informal contacts and visits to scientific conferences (see Table 2.9 in the Appendix for details).

We derive several variables from the questions above. As a first measure, we divide the total number of a firm's interactions by the highest possible value, i.e. seven, to calculate the total interaction intensity ranging between 0 and 1. Next, we calculate the intensities for formal and informal interactions separately. That is, we divide the number of formal interactions by five and the number informal interactions by two. For example, a firm that engages solely in supplier relationships and contract research (both formal), and scientific conferences (informal), is assigned the following interaction intensities: formal 2/5, informal 1/2, and total 3/7. In a similar way, we create intensity variables for occasional and continuous interactions.

Internal R&D is measured as a binary variable that takes the value of one if the firm conducts R&D and zero otherwise. In the sample, 47% of firms stated to conduct own R&D (52% for academic start-ups). The survey did not ask for R&D budgets as founders typically find it difficult to provide such a figure in a telephone interview. Much of the R&D in these new firms is typically conducted in only partially structured routines and by the founder(s) themselves, which makes quantifying R&D expenditure infeasible.

Control variables

A firm's ability to innovate is also influenced by firm characteristics, such as the strength of the firm's resource base and founders' managerial capabilities (Griliches 1992; Love and Roper 1999). Control variables included as a proxy for these dimensions comprise current *firm size*, and information on whether the firm is located in an *innovation center*. In addition, the *founding team size* may contribute to NTBFs' innovation success. Larger founding teams have better resource endowments and abilities to mobilize new competencies, accelerated decision processes (Eisenhardt and Schoonhoven 1990) and have a complementary technical and business skills composition (Colombo and Grilli 2005; Ganotakis 2012; Protogerou et al. 2017). In case of academic start-ups, we control for founders' disciplinary fields of study (*engineering, natural sciences, life sciences, social sciences, business/econ/law, or "other discipline"*) and the presence of a professor in the founding team (*professor in team)*. Finally, we include year dummies to account for cohort effects as well as aggregate sector indicators (*high-tech manufacturing, technology-based services and knowledge-intensive services*). There are significant differences in variable means across all control variables between academic start-ups and non-academic ones (see Table 2.6 in the Appendix).

Model specification

We estimate an innovation production function to measure the influence of knowledge interactions on the probability of introducing market innovations. Since the participation of firms in knowledge exchange with public research institutions does not occur randomly (Bellucci and Pennacchio 2016), we account for this choice to interact with PRIs by estimating selection models for binary outcome variables (Van de Ven and Van Pragg 1981). The selection into knowledge interactions for firm$_i$ is modelled as follows:

$$\Pr(Knowledge\ Interaction)_i = \beta_0 + \sum_{n=1}^{2} \beta\, ER_i + \sum_{n=3}^{k} \beta_n\, C_i + \bar{u}_i \qquad (1)$$

The selection equation (1) models the probability of pursuing knowledge interactions as a function of two exclusion restrictions (*ER*), a vector of control variables *C*, and parameter $\bar{u}$ as the error term.

The model is identified through exclusion restrictions that significantly enter the selection stage but not the outcome stage. First, scientific work experience, that is, the founder had worked or is currently working at a public research institution, provides founders with familiarity of scientific ideas, personnel and knowledge about institutions and resources at the PRIs. Scientific work experience is therefore a strong predictor of direct knowledge interactions. In our sample, scientific work experience is not necessarily directly related to innovation performance. While it is possible to argue that having worked at a PRI is beneficial for innovation performance because of specialized skills and know-how developed during that time, it is also possible to expect the opposite. A lack of industry work experience might reduce the founder's ability to successfully implement market novelties. The data suggests that while there is a 0.17 correlation between this variable (*scientific work experience*) and the innovation indicator, the variable is not significant in the selection stage when we account for other firm and founder characteristics. That is, it should affect innovation mainly through knowledge interactions with which it correlates more strongly and the observed correlation may stem from this indirect link. Secondly, we derive a regional indicator that takes the value one if the start-up is located in a (larger) *city*. A high share of cities in Germany has a higher education institution or public research infrastructure such as a university. Being located closer to a university may increase the likelihood to engage in interactions with researchers in these places. City infrastructure, more generally, may further facilitate both formal and informal knowledge exchange (Glaeser 2007). This variable does not explain innovation performance directly in our sample. In fact, the bilateral correlation with innovation performance is weak (0.01) and *city* does not significantly enter the innovation production function, but is a relatively strong predictor in the selection equation.

First, all models are applied to the full sample. Here, the scientific discipline dummies are included in the selection and the innovation equation to control for academic background, i.e. by adding these variables we control for the specific disciplinary background rather than just accounting for academic education with a single binary indicator. Subsequently, we estimate the models for the two subsamples of academic and non-academic start-ups and the vector of controls is adjusted to the respective sample. For the group of non-academic start-ups in the split sample analysis, we replace scientific work experience with a variable that indicates abandoned studies (*quit studies*) because scientific work experience is rare among non-academic founders. Founders who quit their university education may be less fond of returning to research organizations for collaboration either because they simply do not perceive possible knowledge transfer as useful or because of the negative university experience as such. The correlation between *quit studies* and the dependent variable in the outcome equation is indeed very small (-0.05), whereas it is a significant (negative) predictor of interaction likelihood.

In the outcome equation (2), we estimate a NTBF's probability of introducing a market novelty as a function of *formal* and *informal* interaction intensities derived in the previous section, internal *R&D* and the set of control variables. We account for the selection into knowledge interactions by jointly estimating the innovation equation with the selection equation and taking into account the correlation coefficient ρ (Van de Ven and Van Pragg 1981). A statistically significant $\alpha^{\rho} = 0.5\ln(1+\rho)/(1-\rho)$ indicates that some selection bias would be ignored in the absence of the selection equation.

$$\Pr(Innovation)_i = \beta_0 + \beta_1\, Formal_i + \beta_2 Informal_i + \beta_3 R\&D_i + \sum_{n=5}^{k} \beta_n\, C_i + \alpha^{\rho}{}_i + \tilde{u}_i \qquad (2)$$

Departing from this basic setup, in two further specifications we compare differences between occasional and continuous interactions by estimating separate models for each type. The models which incorporate the occasional interaction variables exclude firms that engage in continuous interaction so that we can compare occasional versus no interaction. In the models with the continuous interaction variables, we leave both no interaction and occasional ones in the comparison group.

Finally, the marginal contribution of each interaction type (i.e. formal/informal) is estimated depending on the simultaneous use of the respective other type as well as depending on the presence of internal R&D activities. We model this by a triple interaction effect between formal interaction intensity, informal interaction intensity and internal R&D:

$$\Pr(Market\ novelty)_i = \beta_0 + \beta_1 Formal_i + \beta_2 Informal_i + \beta_3 R\&D_i + \beta_4 Formal_i Informal_i + \beta_5 Formal_i R\&D_i + \beta_6 Informal_i R\&D_i + \beta_7 Formal_i Informal_i R\&D_i + \sum_{n=8}^{k} \beta_n\, Controls_i + \alpha^{\rho}{}_i + \tilde{u}_i \qquad (3)$$

2.4 Results

Descriptive analysis

Table 2.1 provides first insights on the overall relevance of public research for NTBFs and on the kinds of interactions that they maintain with PRIs. The large majority of the sampled NTBFs maintain some kind of contact with public research institutions whereby informal and occasional interactions with PRIs are more prevalent (83% and 80%) than formal and continuous interactions (79% and 58%). Firms also engage in multiple forms simultaneously with a median number of three different interaction types. Comparing the numbers of the occurrence of a particular aggregate type (any, formal and informal) and the numbers of the individual types further indicate that firms use of several types of interactions simultaneously.

Table 2.1 further shows that student internships are the most common type of formal interaction between NTBFs and PRIs (48%) followed by supplier relationships (40%) and advanced staff training (30%). Joint research activities and contract research are still quite common with 27 and 18%, respectively. Most forms of interactions like staff training, joint research, contract research and visits to scientific conferences are clearly less frequent on a continuous basis while there is no difference for personal contacts. We further observe differences in the use of knowledge interactions between academic start-ups and non-academic start-ups (see Table 2.6 in the Appendix) with the former making use of more different knowledge interactions compared to non-academic start-ups. For example, joint research and contract research with PRIs are performed nearly twice as often in the group of academic start-ups (32 versus 17% and 22 versus 9%). Yet also, informal interactions like personal contacts and conference visits are more common in academic start-ups (64 versus 48% and 80 versus 67%). The only exception are supplier relationships that are more frequent in non-academic start-ups (43 versus 38%).

Table 2.1: Knowledge interactions by interaction persistency

	Combined		Occasional		Continuous	
	mean	std. dev.	mean	std. dev.	mean	std. dev.
Any Interaction	**.92**	**(.27)**	**.80**	**(.40)**	**.58**	**(.49)**
Formal Interactions	**.79**	**(.41)**	**.61**	**(.49)**	**.41**	**(.49)**
Student Internships	.48	(.50)	.26	(.44)	.22	(.41)
Supplier Relationships	.40	(.49)	.23	(.42)	.16	(.37)
Adv. Staff Training	.30	(.46)	.20	(.40)	.10	(.30)
Joint Research	.27	(.44)	.17	(.38)	.10	(.29)
Contract Research	.18	(.38)	.13	(.34)	.04	(.20)
Informal Interactions	**.83**	**(.38)**	**.58**	**(.49)**	**.44**	**(.50)**
Personal Contacts	.76	(.43)	.38	(.48)	.38	(.49)
Scientific Conferences	.59	(.49)	.40	(.49)	.19	(.39)

Notes: Number of observations: 2879. Means and standard deviations of binary variables are shown.

Knowledge interactions, R&D, and the founders' academic backgrounds

Table 2.2 presents sub-sample mean differences in interaction types and intensities between R&D-active and non-R&D-active firms and between academic and non-academic start-ups. We observe that nearly all observed combinations of knowledge interactions (i.e. formal/informal, occasional/continuous, both) are used more often in the group of R&D-active firms with the exception of occasional informal interactions which are likewise prevalent in both groups. The mean difference is largest for those firms which make use of both, formal and informal types of interaction (60 versus 81%). These numbers indicate that conducting own R&D is not a necessary condition for knowledge interactions as also firms without own R&D make use of formal and informal interactions. The share of firms engaging in knowledge interactions is, however, higher in the group of R&D-active firms.

A similar picture emerges when we compare academic start-ups to non-academic start-ups. While both groups are similar likely to interact in any way with PRIs, we find that academic start-ups use more channels (3.2 different types versus 2.4, on average) and maintain these contacts on a more continuous basis. With respect to interaction intensities, we see that the share of firms that use both formal and informal interaction on a continuous basis is twice as high for academic start-ups (7 versus 15% for formal and 18 versus 34% for informal). This finding is not surprising since academic start-ups may naturally be more familiar with higher education institutions and their research facilities and thus are more aware of the potential benefits for their business and technological developments that arise from these sources. Yet, there are no such differences for contacts on an occasional basis.

Table 2.2: Sample mean-comparison by R&D activity and innovation: selected variables

	I. Full Sample	II No internal R&D	III. Internal R&D	II vs. III	III. Non-Academic Start-up	IV. Academic Start-up	III vs. IV
# Observations	2879	1537	1342		921	1958	
	mean (std. dev.)	mean (std. dev.)	mean (std. dev.)	t-test	mean (std. dev.)	mean (std. dev.)	t-test
Knowledge Interactions							
Combined (d)	.92 (.27)	.90 (.30)	.95 (.22)	***	.92 (.27)	.92 (.26)	
Formal interaction (d)	.79 (.41)	.72 (.45)	.87 (.34)	***	.76 (.43)	.80 (.40)	***
Occasional (d)	.61 (.49)	.53 (.50)	.69 (.46)	***	.60 (.49)	.61 (.49)	
Continuous (d)	.41 (.49)	.33 (.47)	.50 (.50)	***	.29 (.46)	.46 (.50)	***
Informal interaction (d)	.83 (.38)	.77 (.42)	.89 (.31)	***	.75 (.43)	.87 (.34)	***
Occasional (d)	.58 (.49)	.59 (.49)	.57 (.50)		.60 (.49)	.57 (.49)	
Continuous (d)	.44 (.50)	.35 (.48)	.54 (.50)	***	.29 (.46)	.51 (.50)	***
Both (d)	.69 (.46)	.60 (.49)	.81 (.39)	***	.59 (.49)	.74 (.44)	***
Interaction Intensities							
Combined	.42 (.25)	.35 (.21)	.51 (.26)	***	.35 (.21)	.46 (.25)	***
Formal interactions	.32 (.26)	.24 (.20)	.42 (.28)	***	.26 (.21)	.35 (.27)	***
Occasional	.20 (.21)	.16 (.18)	.25 (.23)	***	.18 (.19)	.21 (.21)	***
Continuous	.12 (.18)	.08 (.13)	.17 (.22)	***	.07 (.13)	.15 (.20)	***
Informal interactions	.68 (.38)	.61 (.40)	.75 (.34)	***	.57 (.40)	.72 (.36)	***
Occasional	.39 (.38)	.40 (.38)	.38 (.38)		.40 (.37)	.39 (.38)	
Continuous	.29 (.36)	.22 (.32)	.37 (.38)	***	.18 (.30)	.34 (.37)	***

Notes: *** indicates a significance level of 1%; (d) indicates a dummy variable.

Knowledge interactions and innovation performance

Table 2.3 shows the results of the selection equations (*S*) and several specifications of the innovation outcome equation (models 1 - 6). Average marginal effects (AME) in the selection model indicate that internal R&D, firm size and both exclusion restrictions correlate strongly with firms' probability of interacting with PRIs. In model 1, we see that an increase in knowledge interaction intensity by one unit, regardless of its type and intensity, increases the predicted innovation probability by an average of 0.27 percentage points.

In addition to the average effect, we can look at predicted innovation probabilities at different interaction intensity levels. This means, for example that the average predicted innovation probability for a firm without any interaction is 28% while a firm using one interaction type (out of seven) has a 31.7% innovation probably. A firm with three types of interaction has again an about three percentage point higher innovation probably (35%). In the most extreme cases, i.e. for a firm without any contacts to PRIs and a firm that makes use of all seven types, the difference in innovation probably is 28% versus 57%, i.e. 29 percentage points. These marginal effects at representative values (MERs) are significant over the entire range of interaction intensities showing no evidence for decreasing returns to engaging in multiple interaction modes. In model 2, we distinguish between formal and informal interactions and it turns out that an increase of formal interaction intensity by one unit increases the probability of innovation by 0.17 percentage points, while an additional informal interaction is associated with a 0.10 percentage point higher innovation probability, on average. When taking into account interaction persistency in models 3 and 4, this pattern holds for occasional interactions, but not for continuous ones. Continuous formal and informal interaction intensities both have an average marginal impact of about 0.10 percentage point for each additional continuous interaction. However, it should be noted that firms with occasional interaction are in the comparison group, i.e. the marginal innovation advantage needs to be interpreted relative to maintaining less than continuous interaction intensity. Thus for informal types there is a stronger additional benefit of continuous interaction, while for formal types the main benefit stems from having such interactions at all. Formal interactions on a continuous basis still increase innovation probability, but at a slightly smaller marginal rate (0.15 versus 0.10).

These results support *Hypothesis 1* that both formal and informal knowledge interactions are positively associated with innovation outcomes. However, we also see that occasional formal interactions can be impactful, while in the case of informal interaction continuity increases the returns in terms of innovation performance. Thus, *Hypothesis 2* stating that the association between knowledge interactions with PRIs and innovation performance is stronger for continuous interactions than for occasional ones is confirmed for informal interactions, but not for formal ones.

Most control variables in Table 2.3 show the expected relationships with innovation performance. In particular, internal R&D has a strong influence on interaction likelihood as well as on the innovation probability. Firms' size is another significant indicator for contacts to PRIs (in line with Grimpe and Hussinger 2013) and innovation in line with previous literature (Hansen 1992). Moreover, NTBFs residing in an innovation center are more likely to introduce innovations. This can be either due to selection of the most innovative ones into these centers or the supportive environment of innovation incubators and science parks that leverage the innovation performance of start-ups through basic research infrastructure, peer contacts, investors and business know-how (Felsenstein 1994). With respect to sectorial differences, we observe that firms in the high-tech manufacturing sector are more likely to innovate compared to knowledge-intense services while technology-based services are not more likely than knowledge-intense services. The two founding team related variables (professor and founding team size) do not seem to correlate with innovation performance in the presence of the other explanatory variables. Start-ups with larger founding teams, however, are less likely to interact with PRIs. Finally, none of the year or discipline dummies seems to drive innovation performance substantially, except that firms with founders having a background in life sciences or social sciences appear to have lower innovation probabilities than founders with a background in engineering.

Table 2.3: Estimation results from probit models (with selection) on market novelty

	interaction (d)	market novelty (d)			
Model	S	1	2	3	4
		Any	Combined	Occasional	Continuous
Interactions (intensity)		.273***			
		(.042)			
Informal interactions (intensity)			.102***	.066*	.102***
			(.026)	(.038)	(.026)
Formal interactions (intensity)			.173***	.145**	.102**
			(.037)	(.068)	(.052)
R&D (d)	.048***	.297***	.298***	.279***	.312***
	(.010)	(.025)	(.025)	(.040)	(.025)
ln(firm size$_{2008}$)	.046***	.032**	.034***	.018	.041***
	(.007)	(.013)	(.013)	(.022)	(.012)
Innovation centre (d)	.015	.082***	.083***	-.004	.090***
	(.021)	(.032)	(.032)	(.062)	(.032)
Professor in team (d)	.042	.018	.018	.055	.017
	(.037)	(.034)	(.034)	(.105)	(.035)
High-tech manufacturing (d)	-.092***	.053**	.055**	.061	.056**
	(.014)	(.027)	(.026)	(.046)	(.026)
Technology-based services (d)	-.045***	.013	.016	.041	.016
	(.013)	(.023)	(.023)	(.038)	(.023)
Founding team size	-.010**	.003	.003	.022	.002
	(.005)	(.009)	(.008)	(.018)	(.009)
ER 1: City (d)	.021**				
	(.010)				
ER 2: Scientific work experience (d)	.064***				
	(.018)				
Disciplines (d)	yes	yes	yes	yes	yes
Years (d)	yes	yes	yes	yes	yes
Log pseudolikelihood	-2145.350		-2144.840	-1012.255	-2151.623
AIC	4374.699		4375.680	2110.510	4389.246
Wald test of independent equations [chi2(1)]	3.279*		3.411*	4.116**	3.875**
α^{ρ}	-.634*		-.616*	-.717**	-.747**
# observations	2879		2879	1207	2879

Notes: Average marginal effects (AME) presented, standard errors in parenthesis. *** (**,*) indicate a significance level of 1% (5%, 10%); (d) indicate a dummy variable. The group of knowledge-intensive services serves as reference category.

In addition to the full sample models, separate models for academic start-ups and non-academic start-ups are presented in Table 2.4. Regarding the combined knowledge interaction share (models 1 and 4), we find a significant positive effect of additional knowledge interactions which confirms *Hypothesis 1* for both groups, but the AME however is somewhat smaller for non-academic start-ups (0.23 versus 0.28).

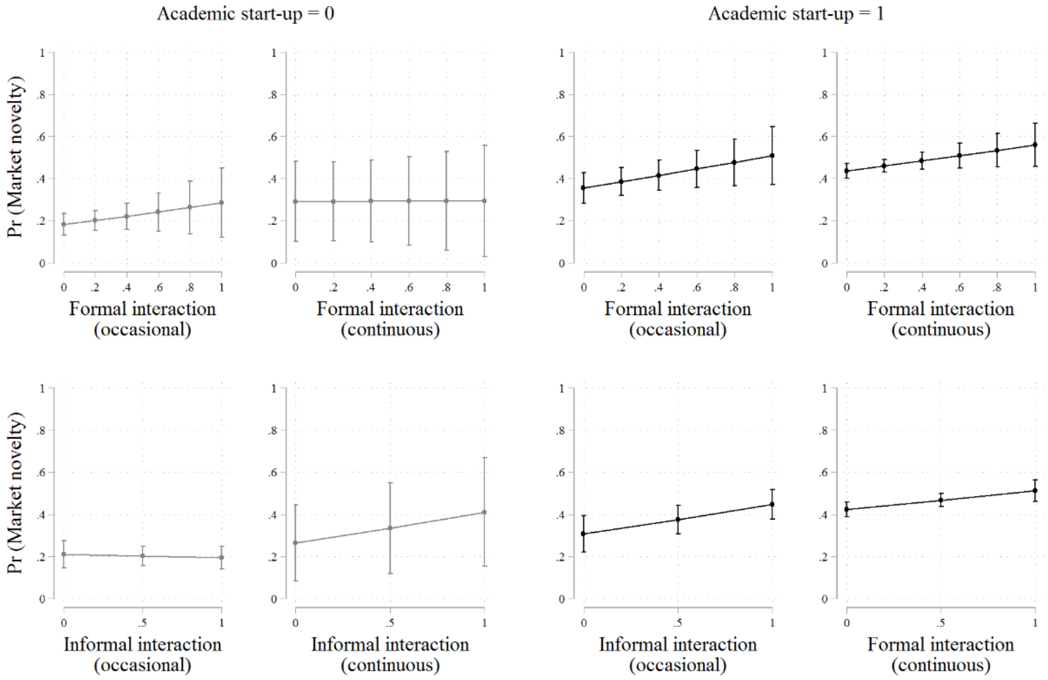
When distinguishing interaction persistency, we see that occasional interactions have little effect on innovation probability for non-academic start-ups (model 2). For academic start-ups, in contrast, we find that each additional occasional interaction of both types increase innovation by 0.14 percentage points for informal and 0.15 percentage points for formal interactions (model 5). Interestingly, innovation performance of non-academic start-ups is influenced by continuous informal interactions with an AME of 0.14 (model 3). For academic start-ups, on the other hand, maintaining continuous interaction of both formal and informal types contributes significantly to innovation performance (model 6). We illustrate the marginal effects at different interaction intensities for academic and non-academic start-ups also graphically. Figure 2.2 depicts the estimated innovation probability at different interaction intensities and persistency levels (see Table 2.10 in the Appendix for the estimated marginal effects at the different levels, MERs). The slope of the predictive margins curve represents the marginal effects between representative values of interaction intensity. The graphically illustration reflects the patterns found in the average marginal effects as reported in Table 2.4. Non-academic start-ups have a higher probability to introduce a market novelty primarily when engaging in more continuous informal interactions types (in line with *Hypothesis 2*). Academic start-ups, in contrast, benefit from additional knowledge interactions regardless of its type. Persistency, however, matters for formal interactions (again in line with *Hypothesis 2*) as the marginal effect of occasional formal interactions is only significant in the first interval, i.e. from zero to one occasional formal interaction.

Table 2.4: Estimation results from probit models (with selection) on market novelty for academic and non-academic start-ups

	Academic start-up = 0				Academic start-up = 1			
	interaction (d)	market novelty (d)			interaction (d)	market novelty (d)		
Model	S1	1	2	3	S2	4	5	6
		any	occasional	continuous		any	occasional	continuous
Interactions (intensity)		.227***				.278***		
		(.068)				(.050)		
Informal interactions (intensity)			-.017	.136**			.139***	.087***
			(.044)	(.059)			(.052)	(.030)
Formal interactions (intensity)			.094	.003			.147*	.121**
			(.093)	(.113)			(.088)	(.058)
R&D (d)	.055***	.305***	.268***	.314***	.048***	.295***	.269***	.317***
	(.017)	(.034)	(.049)	(.066)	(.013)	(.032)	(.057)	(.029)
ln(firm size$_{2008}$)	.060***	.022	.047**	.022	.040***	.048***	.015	.056***
	(.014)	(.016)	(.023)	(.047)	(.008)	(.015)	(.028)	(.014)
Innovation centre (d)	.055	.050	-.046	.053	.020	.085**	.008	.090**
	(.065)	(.069)	(.109)	(.079)	(.022)	(.036)	(.073)	(.036)
Professor in team (d)					.037	.012	.028	.014
					(.035)	(.036)	(.113)	(.036)
High-tech manufacturing (d)	-.004	.004	-.004	.004	-.138***	.084**	.141**	.084**
	(.023)	(.033)	(.046)	(.038)	(.017)	(.037)	(.069)	(.034)
Technology-based services (d)	-.010	-.020	.003	-.013	-.065***	.027	.068	.026
	(.023)	(.036)	(.049)	(.041)	(.016)	(.028)	(.050)	(.027)
Founding team size	-.023	.018	.028	.023	-.010**	.002	.022	.001
	(.018)	(.022)	(.027)	(.032)	(.005)	(.009)	(.022)	(.009)
ER 1: City (d)	.032*				.016			
	(.018)				(.011)			
ER 2: Scientific Work experience (d)					.068***			
					(.018)			
ER 3: Quit studies (d)	-.136**							
	(.069)							
Disciplines (d)	no	no	no	no	yes	yes	yes	yes
Years (d)	yes	yes	yes	yes	yes	yes	yes	yes
Log pseudolikelihood	-667.969		-395.147	-668.811	-1441.772		-582.157	-1447.184
AIC	1391.938		848.294	1395.623	2967.545		1250.315	2980.367
Wald test of independent equations [chi2(1)]	8.494***		.1774	.022	2.752*		4.556**	4.614**
α^{ρ}	2.273***		.426	-.296	-.766*		-1.008**	-.794**
# observations	921		515	921	1958		692	1958

Notes: Average marginal effects presented. *** (**,*) indicate a significance level of 1% (5%, 10%); (d) indicate a dummy variable. The group of knowledge-intense services serves as reference category.

Figure 2.2: Formal and informal interactions and radical innovation probability



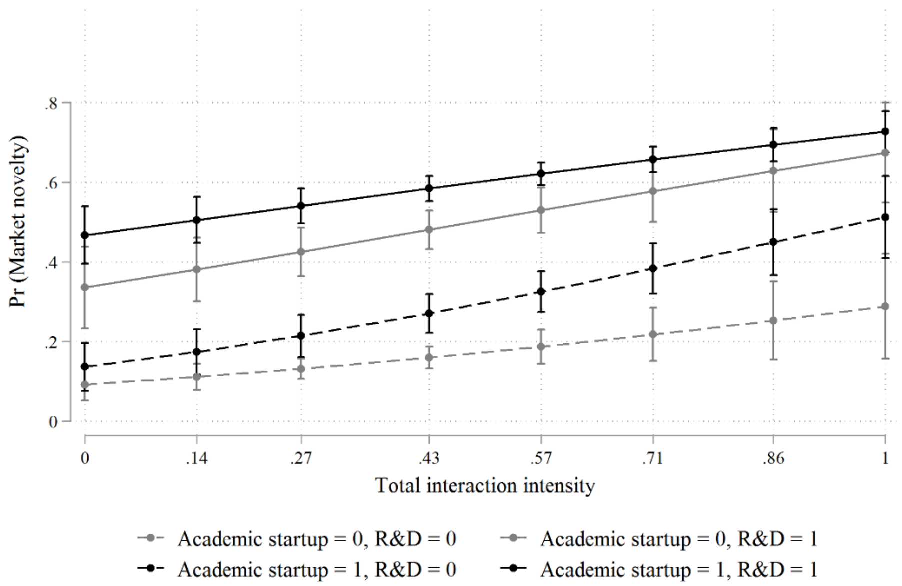
Notes: The graphs show the predicted probability of introducing a market novelty at different values of interaction intensities. Confidence intervals indicate the significance of the predictive margins. For information on the significance of the marginal effects, see Table 2.10 in the Appendix.

The moderating role of internal R&D

To explore the moderating role of R&D, we plot the average predicted innovation probability at different knowledge interactions intensities separately for R&D performers (dashed line) and non-performers (solid line) in Figure 2.3. While R&D performing firms (academic or not) have a higher overall innovation probability, the slope of the curve for R&D-performers is only steeper compared to non-R&D performers in non-academic firms (confirming *Hypothesis 3a*). This is also reflected in the AMEs that are 0.18 for non-R&D performers and 0.34 for firms that engage in internal R&D.

For academic start-ups, the marginal effects in fact increase with interaction intensity for non-R&D performers. In other words, *Hypothesis 3b* is confirmed in the sense that for academic start-ups internal R&D does not necessarily increase the marginal benefit resulting from access to public research. More precisely, for academic start-ups, the AME is 0.37 for non-R&D performers and 0.26 for firms with own R&D. For the latter, the marginal effect increases the more different interaction types are used (see Table 2.11 in the Appendix for the estimated marginal effects).

Figure 2.3: Predictive margins of the total number of interactions on innovation



Notes: The graphs show the predicted probability of introducing a market novelty at different values of interaction intensities. Estimated effects are derived from models as outlines in equation 4 in section 3.2. Confidence intervals indicate the significance of the predictive margins. For information on the significance of the marginal effects, i.e. the slope of the curve, see Table 2.11 in the Appendix.

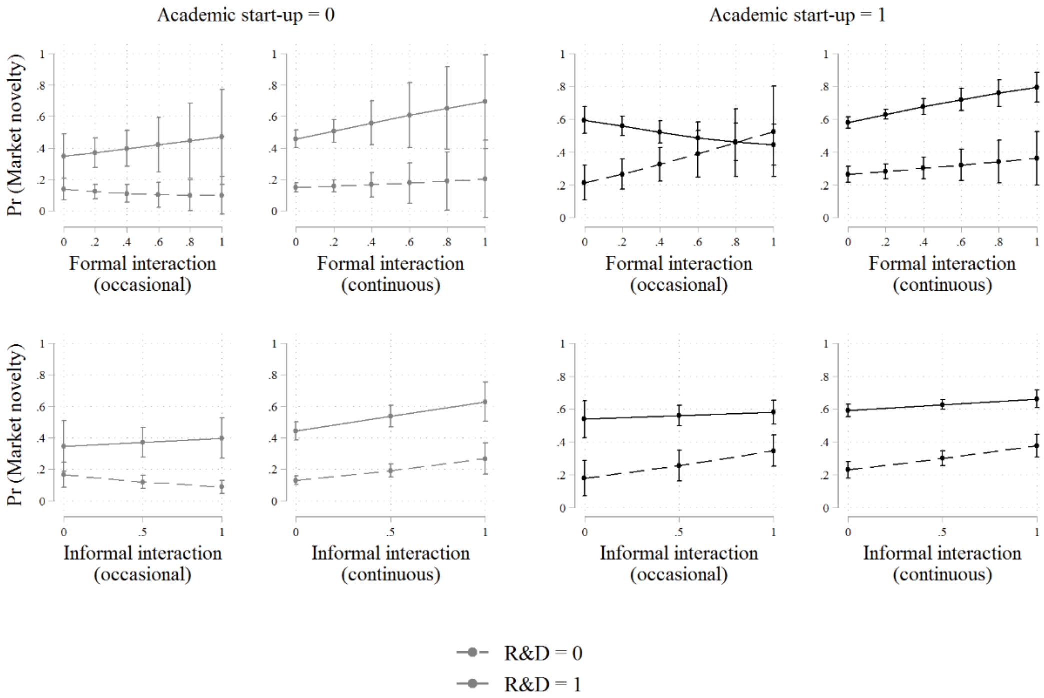
Distinguishing the type of interactions, Figure 2.4 refines the previous results by illustrating that for the group of non-academic start-ups, neither for R&D performing ones nor for others, occasional interactions are significantly related to innovation probability. For continuous informal interactions, the association with innovation probability is stronger for R&D performing firms (AME = 0.18) than for non-R&D performing ones (AME = 0.11). There is also a difference in returns from continuous formal interactions (AME = 0.25 versus 0.04), but here the marginal effects are insignificant over most of the range for both groups (see Table 2.12 in the Appendix for the details).

For academic start-ups and in case of occasional formal interactions, the average marginal increase associated with additional formal interaction types is 0.27 for non-R&D performers and the AME is even negative (-0.16), but not statistically significant, for firms with own R&D. Likewise for occasional informal interactions, there is a positive relationship only for non-R&D performers (AME = 0.17). These results indicate that occasional interactions can provide valuable knowledge in the absence of internal R&D, at least for NTBFs with academic founders. Yet also for continuous informal interactions, the associations is stronger for non-R&D performers (0.14 versus 0.07).

The picture reverses only for continuous formal interactions, where the increase in innovation probability associated with additional continuous formal interactions is stronger when firms also have in-house R&D activities (AME = 0.24 versus AME 0.9). Results presented in Figure 2.4

support *Hypothesis 4* stating that marginal effects on innovation performance are higher if firms engage in internal R&D for non-academic start-ups and in the case of continuous formal interactions also for academic start-ups. This suggests that also academic founders' own R&D activity is complemented by knowledge spillovers from public research when interactions are continuous.

Figure 2.4: Predictive margins of knowledge interactions on innovation



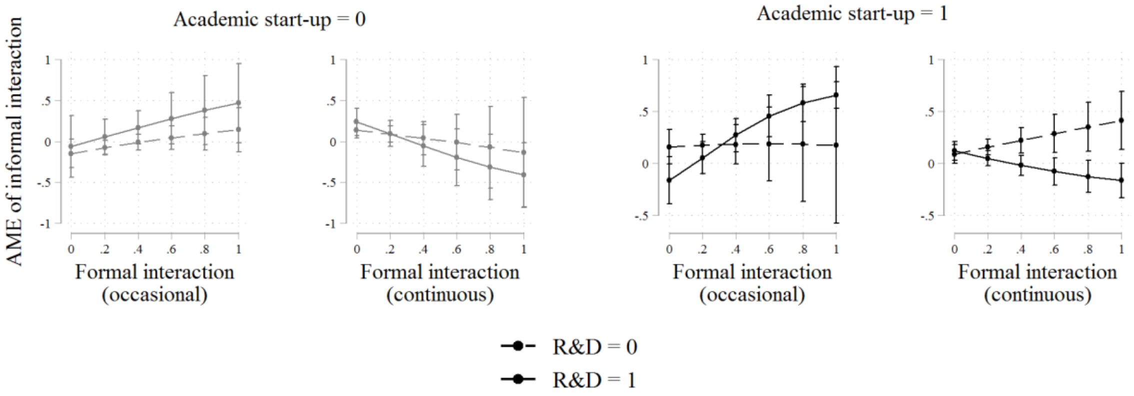
Notes: The graphs show the predicted probability of introducing a market novelty at different values of interaction intensities. Estimated effects are derived from models as outlines in equation 4 in section 3.2. Confidence intervals indicate the significance of the predictive margins. For information on the significance of the marginal effects, i.e. the slope of the curve, see Table 2.12 in the Appendix.

Complementarity between informal and formal interaction

Figure 2.5 shows the results from the test of complementarity between informal and formal interaction modes. For non-academic start-ups, the marginal benefit of occasional informal interactions seems to increase with the simultaneous use of occasional formal interactions independent of own R&D activity. This effect however is never statistically significant (see Table 2.13 in the Appendix). For continuous interactions, the marginal benefit decreases indicating that maintaining both informal and formal interactions on a continuous basis incurs some transaction costs which outweigh possible benefits at least at very high interaction intensities. Yet also here, the marginal effects are not statistically significant except at very low interaction intensities (positive) and very high intensities (negative). Therefore, we cannot clearly conclude that formal and informal interactions are substitutes or complements for non-academic start-ups.

For academic start-ups, but only for R&D-active ones, we see that the marginal effect of occasional informal interactions increases at higher occasional formal interaction intensities. On the contrary, for continuous interactions, firms without own R&D benefit more from continuous informal interactions the more they also engage in formal ones. Thus, for only this group of firms, *Hypothesis 5* of complementarity between formal and informal interactions is confirmed (see Table 2.13 in the Appendix for details on the marginal effects).

Figure 2.5: Average marginal effects of informal interactions at different levels of formal interactions



Notes: The graphs show marginal effect of additional information interactions on the probability of introducing a market novelty at different values of formal interaction intensity. Estimated effects are derived from models as outlines in equation 4 in section 3.2. Confidence intervals indicate the significance of the marginal effects. For detailed information on the magnitude and significance of the marginal effects see Table 2.13 in the Appendix.

2.5 Conclusions and implications

This study built on previous research which stressed the role of knowledge spillovers in entrepreneurial ecosystems (Audretsch 1995) and in particular knowledge spillovers from academic research (Audretsch and Lehmann 2005; Audretsch et al. 2005; Mueller 2006; Audretsch et al. 2011; Guerrero et al. 2015). The presented analysis addressed the interplay between NTBFs' engagement in formal and informal interactions with public research institutions (PRIs) and their innovation potential. For a large sample of NTBFs in Germany, we find that a high proportion of them engages in knowledge interactions at least on an occasional basis stressing the role of PRIs as relevant agents in the entrepreneurial ecosystem. Whereas continuous and formal types of interactions are more common for academic start-ups compared to others, there is little difference in occasional interaction between the groups.

Results from innovation outcome models that account for the selection of firms into such knowledge interactions suggest that overall interactions are related to a higher likelihood of introducing new products and services to the market. These results extend the previously found positive link between direct contacts to public research and innovation performance (Mansfield 1991; 1995; Beise and Stahl 1999; Baum et al. 2000; Cohen et al. 2002; Ganotakis and Love 2012) to the context of new firms in high-tech and knowledge intensive industries. In addition, the results add to previous insights by showing that not only formal interaction (hard commitments) but also informal knowledge exchange (softer commitments) may allow firms in knowledge-intensive industries to benefit from direct links to public research.

The results furthermore confirm previous literature that showed that internal R&D and interaction persistency both moderate the knowledge flows between public research and firm innovation (e.g. for internal R&D, Cohen and Levinthal 1990; Beise and Stahl 1999; Brehm and Lundin 2012; Higón 2016; and for interaction persistency, e.g. Amara and Landry 2005, Nieto and Santamaria 2007; Belderbos et al. 2015). In particular, our results show that the contribution to innovation performance depends not only on the type of interaction, but also on firms' absorptive capacity generated through internal R&D as well as on the persistency with which interactions are maintained. To illustrate the magnitude of the effects, we can look at the substantial differences between firms without interactions and firms that engage heavily in interactions (all different types in our study). For R&D-active, but "non-academic" firms, the difference can add up to 34 percentage points to a firm's probability to introduce a market novelty. For academic start-ups the gain in innovation probability can add up 38 percentage even in absence of own R&D-activity.

Differentiating between formal and informal modes of interaction reveals that while both types are associated with higher innovation performance in academic start-ups, non-academic start-ups mainly benefit from continuous informal interactions. For non-academic start-ups, however, internal R&D is an important prerequisite for benefitting from access to scientific knowledge from public research. This is also the case for academic start-ups, but only for continuous formal interactions. In case of occasional interaction, benefits from interactions with PRIs in terms of innovation are higher for non-R&D performers. This suggests that academic background may substitute for current internal R&D, at least below a certain level of knowledge complexity. Finally, we find some evidence for complementarity between continuous informal and formal interactions for non-R&D performing academic start-ups. This suggests that informal contacts serve translational purposes helping founders to apply research findings to industry uses. Consequently, informal interaction increases in value when combined with formal forms of interaction. Combined use of occasional contacts, on the other hand, is only associated with higher innovation performance in the presence of internal R&D which suggests that for sporadic exchange own knowledge creation matters more.

In spite of efforts to the contrary, the study has some limitations and the presented results need to be interpreted with some caution. Important limitations relate to the level of aggregation at which the interactions are documented. It may be crucial to distinguish between the type of PRI (university, university of applied sciences, or public research organization) and even the research group with which NTBFs engage. Moreover, with the available data, we are unable to track specific relationships between firms and scientists which would enable the studying of relationship-based associations like interaction initiation, interaction duration, content, or remuneration (Goel et al. 2017). Moreover, departing from the linear model of science and technology which sees academic science to be unilaterally shaping and supporting industrial innovation (Kline and Rosenberg 1986), reverse knowledge spillovers or industry-to-academia feedbacks have the potential to inspire public researchers' agendas with practical needs and future applications (Sohn 2014), but could not be incorporated in our analysis.

More generally, the cross-sectional nature of the data does not permit analysis over time. It would be highly desirable to track firms and their use of knowledge interactions over time. Another limitation stems from the relatively short life span of the surveyed firms. In-house basic research or external knowledge sourcing in high and medium-high technology sectors may take up to five years to materialize in innovation (Higón 2016). Our time frame between foundation and self-report is at maximum 4 years and some investments in internal R&D or benefits from publicly funded research may not have become visible yet. We therefore encourage future work on the evaluation of the impact of knowledge interactions over a longer time horizon to derive conclusions regarding the use and impact of public research as firms mature.

Our findings have implications for founders as well as for science, innovation and entrepreneurship policy. Founders in knowledge-intense industries may increase their innovation performance if they engage in multiple formal and informal knowledge interactions with PRIs. Even when they are faced with resource constraints or other factors that prevent them from performing R&D internally, it is still beneficial to regularly tap advice from and exchange with domain experts. The conclusions from this study are relevant for policy makers given the economic impact of universities and public research organizations beyond academic outcomes. We confirm their role as relevant source of knowledge spillovers and as valuable partners for collaboration in the entrepreneurial ecosystem. These insights further confirm the importance of public research funding as policy tool for fostering technology-based entrepreneurship. While policy makers have little direct influence on large corporations as a source of spillovers, they can influence universities and public research organizations as a promoter for entrepreneurship. Policy instruments could also be designed such that they facilitate continuous knowledge transfer between start-ups and public research institutions more explicitly. It seems furthermore advisable to expand existing programs that include knowledge brokering services and assistance in (cooperation) partner search, support of public-private joint projects, and co-location of complementary actors to facilitate informal interaction as means to improve new firms' innovativeness.

References

Acs, Z. J., Autio, E., and L. Szerb (2014). National systems of entrepreneurship: Measurement issues and policy implications. *Research Policy*. 43(3), 476-494.

Almus, M. and E. A. Nerlinger (1999). Growth of new technology-based firms: Which factors matter?. *Small Business Economics*, 13(2), 141-154.

Amara, N., and R. Landry (2005). Sources of information as determinants of novelty of innovation in manufacturing firms: Evidence from the 1999 statistics Canada Innovation Survey. *Technovation*, 25(3), 245–259.

Antonelli, C. (1999). The evolution of the industrial organisation of the production of knowledge. *Cambridge Journal of Economics*, 23(2), 243–260.

Audretsch, D. B. (1995). Innovation, growth and survival. *International journal of industrial organization*, 13(4), 441-457.

Audretsch, D. B., and E. E. Lehmann (2005). Does the knowledge spillover theory of entrepreneurship hold for regions?. *Research Policy*, 34(8), 1191-1202.

Audretsch, D. B., Lehmann, E. E., and S. Warning (2005). University spillovers and new firm location. *Research Policy*, 34(7), 1113-1122.

Autio, E., Kenney, M., Mustar, P., Siegel, D., and M. Wright (2014). Entrepreneurial innovation: The importance of context. *Research Policy*, 43(7), 1097-1108.

Autio, E., and H. Yli-Renko (1998). New, technology-based firms in small open economies - an analysis based on the Finnish experience. *Research Policy*, 26(9), 973-987.

Baum, J. A., Calabrese, T., and B. S. Silverman (2000). Don't go it alone: Alliance network composition and startups' performance in Canadian biotechnology. *Strategic Management Journal*, 21(3), 267-294.

Beise, M., and H. Stahl (1999). Public research and industrial innovations in Germany. *Research Policy*, 28(4), 397-422.

Belderbos, R., Carree, M., Lokshin, B., and J. Fernández Sastre (2015). Inter-temporal patterns of R&D collaboration and innovative performance. *Journal of Technology Transfer*, 40(1), 123–137.

Bellucci, A., and L. Pennacchio (2016). University knowledge and firm innovation: Evidence from European countries. *The Journal of Technology Transfer*, 41(4), 730-752.

Brehm, S., and N. Lundin (2012). University-industry linkages and absorptive capacity: An empirical analysis of China's manufacturing industry. *Economics of Innovation and New Technology*, 21(8), 837-852.

Callon, M. (1994). Is science a public good?. *Science, Technology and Human Values*, 19(4), 395-424.

Catozzella, A., and M. Vivarelli (2014). The catalysing role of in-house R&D in fostering complementarity among innovative inputs. *Industry and Innovation*, 21(3), 179-196.

Cohen, W. M., and D. A. Levinthal (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1), 128-152.

Cohen, W. M., Nelson, R. R., and J. P. Walsh (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science*, 48(1), 1–23.

Colombo, M. G., and L. Grilli (2005). Founders' human capital and the growth of new technology-based firms: A competence-based view. *Research Policy*, 34(6), 795–816.

Dasgupta, P. and P. David (1994). Towards a new economics of science. *Research Policy*, 23(5), 487-521.

Eisenhardt, K. M., and C. B. Schoonhoven (1990). Organizational growth: Linking founding team, strategy, environment, and growth among U.S. semiconductor ventures 1978-1988. *Administrative Science Quarterly*, 35(3), 504.

Felsenstein, D. (1994). University-related science parks–seedbeds or enclaves of innovation?. *Technovation*, 14(2), 93–110.

Fleming, L., and O. Sorenson (2004). Science as a map in technological search. *Strategic Management Journal*, 25(8-9), 909–928.

Fryges, H., Gottschalk, S., and K. Kohn (2010). The KFW/ZEW start-up Panel: Design and research potential. *Schmollers Jahrbuch*, 130(1), 117-131.

Gambardella, A. (1992). Competitive advantages from in-house scientific research: The US pharmaceutical industry in the 1980s. *Research Policy*, 21(5), 391–407.

Ganotakis, P., and J. H. Love (2012). The innovation value chain in new technology-based firms: Evidence from the U. K.. *Journal of Product Innovation Management*, 29(5), 839-860.

Ganotakis, P. (2012). Founders' human capital and the performance of UK new technology-based firms. *Small Business Economics*, 39(2), 495–515.

Glaeser, E. L. (2007). Entrepreneurship and the city. *National Bureau of Economic Research*, Working Paper No. 13551.

Godin, B., and Y. Gingras (2000). The place of universities in the system of knowledge production. *Research Policy*, 29(2), 273-278.

Godoe, H. (2000). Innovation regimes, R&D and radical innovations in telecommunications. *Research Policy*, 29(9), 1033–1046.

Goel, R. K., Göktepe-Hultén, D., and C. Grimpe (2017). Who instigates university–industry collaborations? University scientists versus firm employees. *Small Business Economics*, 48(3), 503–524.

Guerrero, M., Cunningham, J. A., and D. Urbano (2015). Economic impact of entrepreneurial universities' activities: An exploratory study of the United Kingdom. *Research Policy*, 44(3), 748-764.

Griliches, Z. (1992). The search for R&D spillovers. *Scandinavian Journal of Economics*, 94, 29–47.

Grimpe, C., and K. Hussinger (2013). Formal and informal knowledge and technology transfer from academia to industry: Complementarity effects and innovation performance. *Industry and Innovation*, 20(8), 683-700.

Grimpe, C., and W. Sofka (2016). Complementarities in the search for innovation - managing markets and relationships. *Research Policy*, 45(10), 2036-2053.

Gulati, R. (1995). Social structure and alliance formation patterns: A longitudinal analysis. *Administrative Science Quarterly*, 40(4), 619–652.

Hansen, J. A. (1992). Innovation, firm size, and firm age. *Small Business Economics*, 4(1), 37-44.

Higón, D. A. (2016). In-house versus external basic research and first-to-market innovations. *Research Policy*, 45(4), 816-829.

Hottenrott, H., and C. Lawson (2014). Research grants, sources of ideas and the effects on academic research. *Economics of Innovation and New Technology*, 23(2), 109–133.

Klevorick, A. K., Levin, R., Nelson, R., and S. Winter (1995). On the sources and significance of inter-industry differences in technological opportunities. *Research Policy*, 24(2), 185–205.

Kline, S.J., and N. Rosenberg (1986). An overview of innovation. In R. Landau and N. Rosenberg (eds.), *The positive sum strategy: Harnessing technology for economic growth*. Washington, D.C.: National Academy Press, pp. 275–305.

Lockett, A., Siegel, D., Wright, M., and M. D. Ensley (2005). The creation of spin-off firms at public research institutions: Managerial and Policy Implications. *Research policy*, 34(7), 981-993.

Love, J. H., and S. Roper (1999). The determinants of innovation: R&D, technology transfer and networking effects. *Review of Industrial Organization*, 15(1), 43–64.

Lucena, A., and S. Roper (2016). Absorptive capacity and ambidexterity in R&D: Linking technology alliance diversity and firm innovation. *European Management Review*, 13(3), 159-178.

Lynskey, M. J. (2009). Knowledge spillovers from public research institutions to the private sector: Evidence from Japanese new technology-based firms. *International Journal of Technology Transfer and Commercialisation*, 8(2-3), 159-184.

Mansfield, E. (1991). Academic research and industrial innovation. *Research Policy*, 20(1), 1-12.

Mansfield, E. (1995). Academic research underlying industrial innovations: Sources, characteristics, and financing. *The Review of Economics and Statistics*, 77(1), 55-65.

Martin, S., and J. T. Scott (2000). The nature of innovation market failure and the design of public support for private innovation. *Research Policy*, 29(4), 437-447.

Mason, C., and R. Brown (2014). Entrepreneurial ecosystems and growth oriented entrepreneurship, *Final Report to OECD, Paris*, 30(1), 77-102.

Meyer-Krahmer, F. and U. Schmoch (1998). Science-based technologies: University–industry interactions in four fields. *Research Policy*, 27(8), 835-851.

Mueller, P. (2006). Exploring the knowledge filter: How entrepreneurship and university–industry relationships drive economic growth. *Research policy*, 35(10), 1499-1508.

Murray, F. (2004). The role of academic inventors in entrepreneurial firms: Sharing the laboratory life. *Research Policy*, 33(4), 643–659.

Nelson, R. R. (1982). The role of knowledge in R&D efficiency. *Quarterly Journal of Economics*, 97 (3), 453–470.

Nieto, M. J., and L. Santamaría (2007). The importance of diverse collaborative networks for the novelty of product innovation. *Technovation*, 27(6), 367-377.

OECD (2015). *OECD Science, Technology and Industry Scoreboard 2015: Innovation for growth and society.* OECD Publishing, Paris.

OECD (2016). *Main Science and Technology Indicators, Volume 2016 Issue 1*, OECD Publishing, Paris.

Olmos-Peñuela, J., Molas-Gallart, J., and E. Castro-Martinez (2014). Informal collaborations between social sciences and humanities researchers and non-academic partners'. *Science and Public Policy*, 41(4), 1-14.

Perkmann, M., and K. Walsh (2009). The two faces of collaboration: Impacts of university-industry relations on public research. *Industrial and Corporate Change*, 18(6), 1033-1065.

Ponomariov, B., and P. C. Boardman (2008). The effect of informal industry contacts on the time university scientists allocate to collaborative research with industry. *Journal of Technology Transfer*, 33(3), 301–313.

Powell, W. W., Koput, K. W., Smith-Doerr, L., and J. Owen-Smith (1999). Network position and firm performance: Organizational returns to collaboration in the biotechnology industry. *Research in the Sociology of Organizations*, 16(1), 129-159.

Protogerou, A., Caloghirou, Y., and N. S. Vonortas (2017). Determinants of young firms' innovative performance: Empirical evidence from Europe. *Research Policy*, 46(7), 1312-1326

Qian, H., and Z. J. Acs (2013). An absorptive capacity theory of knowledge spillover entrepreneurship. *Small Business Economics*, 40(2), 185-197.

Roper, S., Love, J. H., and K. Bonner (2017). Firms' knowledge search and local knowledge externalities in innovation performance. *Research Policy*, 46(1), 43-56.

Siegel, D. S., Waldman, D. A., Atwater, L. E., and A. N. Link (2004). Toward a model of the effective transfer of scientific knowledge from academicians to practitioners: Qualitative evidence from the commercialization of university technologies. *Journal of Engineering and Technology Management*, 21(1), 115-142.

Sohn, E. (2014). The endogeneity of academic science to local industrial R&D. *Academy of Management Proceedings*, 14(1), 11413).

Stinchcombe, A. L. (1965). Social structure and organizations. In J. G. March (ed.), *Handbook of Organizations*, Rand McNally, Chicago, IL, pp. 142–193.

Storey, D. J., and B. S. Tether (1998). New technology-based firms in the European Union: an introduction. *Research Policy*, 26(9), 933-946.

Tether, B. S. (2002). Who co-operates for innovation, and why: an empirical analysis. *Research Policy*, 31(6), 947-967.

Van de Ven, W. P. M. M., and B. M. S. Van Praag (1981). The demand for deductibles in private health insurance. *Journal of Econometrics*, 17(2), 229–252.

Vega-Jurado, J., Gutiérrez-Gracia, A., and I. Fernández-de-Lucio (2009). Does external knowledge sourcing matter for innovation? Evidence from the Spanish manufacturing industry. *Industrial and Corporate Change*, 18(4), 637-670.

Zucker, L. G., Darby, M. R., and J. Armstrong (1998). Geographically localized knowledge: Spillovers or markets. *Economic Inquiry* 36(1), 65–86.

Appendix

Table 2.5: Variable definitions and summary statistics

	Min.	Max.	Median	Mean	s.d.	Description
Market novelty	0	1	0	0.37	0.48	Firm introduced a market innovation
# of different interactions	0	7	3	2.97	1.72	Number of different interactions
A: Joint research	0	1	0	0.27	0.44	Firm performed joint research with PRI
B: Science as customer	0	1	0	0.40	0.49	Firm supplied/equipped PRIs
C: Contract research	0	1	0	0.18	0.38	Firm assigned research contacts to university
D: Staff training	0	1	0	0.30	0.46	Firm educated its staff at the university
E: Student internships	0	1	0	0.48	0.50	Firm provided student internships / theses writing
F: Scientific conferences	0	1	1	0.59	0.49	Firm visited scientific conferences
G: Personal contacts	0	1	1	0.76	0.43	Firm maintained informal personal contacts to university
Total interaction	0	1	0	0.42	0.25	# of total interaction as share of all total [intensity]
Formal interaction	0	1	0	0.32	0.26	# of formal interaction as share of all formal [intensity]
Informal interaction	0	1	1	0.68	0.38	# of informal interaction as share of all informal [intensity]
R&D	0	1	0	0.47	0.50	Firm performs internal R&D
Ln(employees)	0.69	5.99	1.97	1.90	0.88	Logged number of employees in 2008
Academic start-up	0	1	1	0.68	0.47	Firm has at least one academic founding team member
Innovation center	0	1	0	0.08	0.27	Firm was located at an innovation center
Professor in team	0	1	0	0.07	0.26	Firm has at least one professor in the founding team
Founding team size	1	14	1	1.67	1.15	Number of founding team members
Engineering	0	1	0	0.26	0.44	At least one founder with engineering background
Natural sciences	0	1	0	0.16	0.37	At least one founder with natural sciences background
Life sciences	0	1	0	0.04	0.20	At least one founder with life sciences background
Social sciences	0	1	0	0.08	0.27	At least one founder with social sciences background
Business/econ/law	0	1	0	0.23	0.42	At least one founder with business, econ, law background
Other discipline	0	1	0	0.03	0.18	At least one founder with other academic background
Quit studies	0	1	0	0.06	0.23	At least one founder has quit higher education studies
City	0	1	1	0.63	0.48	Firm is located in a city
Scientific work experience	0	1	0	0.20	0.40	At least one founder with scientific work experience
High-tech manufacturing	0	1	0	0.33	0.47	Sector 1: Firm operates in high-tech manufacturing
Technology-based services	0	1	0	0.33	0.47	Sector 2: Firm operates in high-tech services
Knowledge-intensive	0	1	0	0.34	0.47	Sector 3: Firm operates in knowledge-intense services
Year 2001	0	1	0	0.16	0.36	Firm founded in 2001
Year 2002	0	1	0	0.17	0.37	Firm founded in 2002
Year 2003	0	1	0	0.18	0.38	Firm founded in 2003
Year 2004	0	1	0	0.16	0.37	Firm founded in 2004
Year 2005	0	1	0	0.19	0.39	Firm founded in 2005
Year 2006	0	1	0	0.14	0.35	Firm founded in 2006

Table 2.6: Variable mean comparisons for main variables by academic background

Variable	Full sample	Academic start-up = 0	Academic start-up = 1	t-test
	mean (standard deviation)			
Market novelty	.37 (.48)	.28 (.45)	.42 (.49)	***
# of different interactions	2.97 (1.72)	2.44 (1.48)	3.22 (1.77)	***
A: Joint research	.27 (.44)	.17 (.38)	.32 (.46)	***
B: Science as customer	.40 (.49)	.43 (.50)	.38 (.49)	***
C: Contract research	.18 (.38)	.09 (.29)	.22 (.41)	***
D: Staff training	.30 (.46)	.23 (.42)	.34 (.47)	***
E: Student internships	.48 (.50)	.36 (.48)	.54 (.50)	***
F: Scientific conferences	.59 (.49)	.48 (.50)	.64 (.48)	***
G: Personal contacts	.76 (.43)	.67 (.47)	.80 (.40)	***
Total interaction [intensity]	.42 (.25)	.35 (.21)	.46 (.25)	***
Formal interaction [intensity]	.32 (.26)	.26 (.21)	.36 (.27)	***
Informal interaction [intensity]	.68 (.38)	.57 (.4)	.72 (.36)	***
R&D	.47 (.50)	.35 (.48)	.52 (.50)	***
Employees	10.04 (20.41)	8.78 (17.68)	10.64 (21.55)	**
Innovation center	.08 (.27)	.03 (.18)	.10 (.30)	***
City	.63 (.48)	.57 (.50)	.65 (.48)	***
Founding team size	1.67 (1.15)	1.22 (.57)	1.89 (1.28)	***
High-tech manufacturing	.33 (.47)	.43 (.49)	.29 (.45)	***
Technology-based services	.33 (.47)	.31 (.46)	.33 (.47)	
Knowledge-intense services	.34 (.47)	.26 (.44)	.38 (.49)	***
Observations	N =2879	N = 921	N = 1958	
	100%	31.99%	68.01%	

Note: t-test for differences in variable means.

Table 2.7: Correlation matrix of the main variables

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 Market novelty	1															
2 Total interaction	.29	1														
3 Informal interaction	.18	.71	1													
4 Formal interaction	.29	.91	.36	1												
5 R&D (d)	.42	.34	.18	.35	1											
6 Firm size	.10	.15	.08	.16	.08	1										
7 Innovation center (d)	.14	.17	.08	.18	.15	.02	1									
8 Academic start-up (d)	.13	.21	.18	.17	.17	.04	.11	1								
9 Professor in team (d)	.13	.26	.15	.26	.18	.06	.13	.19	1							
10 High-tech manufacturing (d)	.08	-.07	-.12	-.02	.20	.04	.00	-.14	-.04	1						
11 Technology-based services (d)	-.01	.01	-.05	.04	-.02	-.06	.05	.02	.05	-.49	1					
12 Knowledge-intense services (d)	-.07	.06	.17	-.02	-.18	.02	-.05	.12	-.01	-.51	-.50	1				
13 Founding team size (d)	.13	.16	.09	.16	.17	.13	.12	.27	.31	.02	.02	-.04	1			
14 City (d)	.01	.05	.04	.05	-.02	.01	.01	.08	.04	-.21	.12	.10	.04	1		
15 Scientific work experience (d)	.17	.30	.23	.27	.21	.03	.13	.35	.55	-.07	.06	.01	.24	.08	1	
16 Quit studies (d)	-.05	-.09	-.10	-.06	-.02	.02	-.02	.15	-.07	-.03	.05	-.02	-.07	-.01	-.08	1

Note: (d) indicates a dummy variable.

Table 2.8: Correlation matrix of knowledge interactions

	A	B	C	D	E	F	G
A: Joint research	1						
B: Science as customer	.17***	1					
C: Contract research	.45***	.10***	1				
D: Staff training	.22***	.07***	.18***	1			
E: Student internships	.22***	.05**	.19***	.25***	1		
F: Scientific conferences	.30***	.08***	.21***	.31***	.19***	1	
G: Personal contacts	.36***	.17***	.26***	.23***	.25***	.39***	1

Notes: *** (**) indicate a significance level of 1% (5%). Number of observations: 2879.

Table 2.9: Factor analysis of knowledge interactions

Variable	Factor 1	Factor2	Uniqueness
A: Joint research	**.6619**	-.2911	.4772
B: Science as customer	**.1539**	-.0772	.9703
C: Contract research	**.5192**	-.2379	.6739
D: Staff training	**.3799**	.1438	.8350
E: Student internships	**.3271**	.0867	.8855
F: Scientific conferences	.5104	**.3959**	.5828
G: Informal contacts	.4765	**.2338**	.7182

Notes: Factor loadings and unique variances presented. Number of observations: 2879.

Table 2.10: Average marginal effects of formal and informal interactions on innovation (corresponds to Figure 2.2)

KI type	KI persistency	KI intensity	dy/dx		Std. err.	dy/dx		Std. err.
			Academic start-up = 0			Academic start-up = 1		
Formal	occasional	0	.089		.083	.145	*	.086
Formal	occasional	.2	.095		.094	.149		.091
Formal	occasional	.4	.100		.105	.152		.094
Formal	occasional	.6	.106		.115	.154		.096
Formal	occasional	.8	.111		.124	.155		.097
Formal	occasional	1	.116		.133	.156		.095
Informal	occasional	0	-.018		.046	.130	***	.044
Informal	occasional	.5	-.017		.044	.140	***	.053
Informal	occasional	1	-.017		.042	.146	**	.058
Formal	continuous	0	.003		.113	.122	**	.059
Formal	continuous	.2	.003		.113	.123	**	.060
Formal	continuous	.4	.003		.113	.124	**	.060
Formal	continuous	.6	.003		.113	.124	**	.060
Formal	continuous	.8	.003		.114	.123	**	.059
Formal	continuous	1	.003		.114	.123	**	.057
Informal	continuous	0	.132	**	.058	.088	***	.030
Informal	continuous	.5	.148	**	.065	.089	***	.031
Informal	continuous	1	.158	**	.067	.089	***	.031

Notes: *** (**,*) indicate a significance level of 1% (5%, 10%). KI stands for knowledge interaction.

Table 2.11: Average marginal effects of formal and informal interaction on innovation (corresponds to Figure 2.3)

KI type	KI persistency	KI intensity	Internal R&D	dy/dx		Std. err.	dy/dx		Std. err.
				Academic start-up = 0			Academic start-up = 1		
formal and informal	combined	0	0	.127	***	.039	.247	***	.032
formal and informal	combined	.143	0	.146	***	.055	.289	***	.048
formal and informal	combined	.286	0	.164	**	.071	.328	***	.066
formal and informal	combined	.429	0	.187	**	.092	.373	***	.088
formal and informal	combined	.571	0	.207	*	.110	.405	***	.103
formal and informal	combined	.714	0	.226	*	.127	.430	***	.112
formal and informal	combined	.857	0	.246	*	.143	.446	***	.115
formal and informal	combined	1	0	.262	*	.155	.449	***	.108
formal and informal	combined	0	1	.319	***	.108	.273	***	.071
formal and informal	combined	.143	1	.334	***	.124	.274	***	.074
formal and informal	combined	.286	1	.343	**	.135	.272	***	.074
formal and informal	combined	.429	1	.349	**	.141	.268	***	.073
formal and informal	combined	.571	1	.348	**	.139	.261	***	.070
formal and informal	combined	.714	1	.343	***	.132	.252	***	.065
formal and informal	combined	.857	1	.331	***	.118	.241	***	.058
formal and informal	combined	1	1	.316	***	.099	.228	***	.050

Notes: *** (**,*) indicate a significance level of 1% (5%, 10%). KI stands for knowledge interaction.

Table 2.12: Average marginal effects of formal interactions on innovation (corresponds to Figure 2.4)

KI type	KI persistency	KI intensity	Internal R&D	dy/dx	Std. err.	dy/dx	Std. err.
				Academic start-up = 0		Academic start-up = 1	
formal	occasional	0	0	-.092	.168	.249 **	.114
formal	occasional	.2	0	-.071	.132	.282 *	.159
formal	occasional	.4	0	-.050	.101	.310	.199
formal	occasional	.6	0	-.030	.081	.331	.227
formal	occasional	.8	0	-.013	.077	.343	.237
formal	occasional	1	0	.002	.082	.345	.226
formal	occasional	0	1	.109	.266	-.150	.127
formal	occasional	.2	1	.121	.262	-.185	.159
formal	occasional	.4	1	.128	.253	-.186	.152
formal	occasional	.6	1	.129	.237	-.154	.109
formal	occasional	.8	1	.123	.217	-.106	.069
formal	occasional	1	1	.113	.195	-.060	.053
formal	continuous	0	0	.035	.120	.089	.107
formal	continuous	.2	0	.042	.134	.095	.110
formal	continuous	.4	0	.049	.149	.098	.112
formal	continuous	.6	0	.056	.165	.100	.112
formal	continuous	.8	0	.063	.182	.099	.110
formal	continuous	1	0	.070	.197	.096	.107
formal	continuous	0	1	.250	.236	.249 ***	.096
formal	continuous	.2	1	.255	.245	.241 ***	.090
formal	continuous	.4	1	.251	.230	.227 ***	.078
formal	continuous	.6	1	.237	.195	.208 ***	.061
formal	continuous	.8	1	.216	.146	.185 ***	.044
formal	continuous	1	1	.190 **	.092	.161 ***	.030
informal	occasional	0	0	-.104	.083	.136 ***	.040
informal	occasional	.5	0	-.076	.053	.169 ***	.063
informal	occasional	1	0	-.051 *	.031	.195 **	.085
informal	occasional	0	1	.046	.138	.033	.080
informal	occasional	.5	1	.051	.137	.046	.092
informal	occasional	1	1	.052	.133	.032	.095
informal	continuous	0	0	.104 ***	.040	.129 ***	.036
informal	continuous	.5	0	.138 **	.066	.149 ***	.048
informal	continuous	1	0	.169 *	.088	.158 ***	.054
informal	continuous	0	1	.184 **	.091	.073	.045
informal	continuous	.5	1	.187 **	.093	.071	.044
informal	continuous	1	1	.175 **	.076	.067 *	.041

Notes: *** (**,*) indicate a significance level of 1% (5%, 10%). KI stands for knowledge interaction.

Table 2.13: Average marginal effects of informal interactions on innovation at different levels of formal interaction intensity (corresponds to Figure 2.5)

KI type	KI persistency	Formal KI intensity	Internal R&D	dy/dx		Std. err.	dy/dx		Std. err.
				Academic start-up = 0			Academic start-up = 1		
informal	occasional	0	0	-.148		.105	.160		.100
informal	occasional	.2	0	-.076		.052	.174	***	.066
informal	occasional	.4	0	-.012		.057	.183		.114
informal	occasional	.6	0	.044		.088	.187		.216
informal	occasional	.8	0	.094		.122	.185		.336
informal	occasional	1	0	.141		.162	.177		.458
informal	occasional	0	1	-.061		.230	-.162		.137
informal	occasional	.2	1	.053		.134	.054		.094
informal	occasional	.4	1	.169		.123	.273	***	.098
informal	occasional	.6	1	.280		.191	.457	***	.121
informal	occasional	.8	1	.381		.256	.583	***	.110
informal	occasional	1	1	.467		.294	.656	***	.078
informal	continuous	0	0	.134	**	.058	.091	*	.054
informal	continuous	.2	0	.090		.061	.156	***	.046
informal	continuous	.4	0	.040		.124	.222	***	.074
informal	continuous	.6	0	-.013		.207	.288	***	.111
informal	continuous	.8	0	-.071		.302	.352	**	.144
informal	continuous	1	0	-.133		.407	.412	**	.170
informal	continuous	0	1	.237	**	.103	.120	**	.055
informal	continuous	.2	1	.095		.097	.049		.045
informal	continuous	.4	1	-.052		.154	-.018		.059
informal	continuous	.6	1	-.192		.211	-.077		.079
informal	continuous	.8	1	-.313		.240	-.126		.093
informal	continuous	1	1	-.408	*	.238	-.165		.101

Notes: *** (**,*) indicate a significance level of 1% (5%, 10%). KI stands for knowledge interaction.

3. Effects of public and private sector consulting on academic research

Abstract*

Academic consulting is an important and effective means of knowledge transfer between the public and private sectors. It offers opportunities for research application but also raises concerns over potentially negative consequences for academic research and its dissemination. For a sample of social, natural and engineering science academics in Germany, and controlling for the selection into consulting, we investigate the effect of consulting with public and private sector organizations on research performance. While previous research suggested that consulting activities might come at the cost of reduced research output, our analysis provides a more nuanced picture. Public sector consulting comes with lower average citations, particularly for junior researchers. Moreover, engagement in consulting increases the probability to cease publishing research altogether, particularly for private sector consulting. The probability of exit from academic research increases with the intensity of consulting engagement for those at the start or towards the end of their academic career and in fields for which the public-private wage gap and opportunities for engagement in duties outside academia are higher. We draw lessons for research institutions and policy about the promotion of academic consulting.

Acknowledgements

The authors thank the ZEW for providing the survey data and particularly Maikel Pellens for his support. The article benefitted from comments and suggestion made by participants of the workshop “Economics of Entrepreneurship and Innovation” at the University of Trier, the ZEW-INNOPAT conference 2015 in Mannheim, the DRUID15 conference in Rome, the EARIE2015 conference in Munich, Technology Transfer Society Annual Conference 2015 in Dublin, and the INCHER seminar at the University of Kassel. The authors would also like to thank two anonymous reviewers for very helpful comments. This research was funded by the DFG research grant no. HO-5390/1-1

* This chapter is based on the peer-reviewed publication “What’s the Price of Academic Consulting? Effects of Public and Private Sector Consulting on Academic Research”, co-authored by Hanna Hottenrott and Cornelia Lawson, forthcoming in *Industrial and Corporate Change*.

3.1 Introduction

In recent years, universities have become more proactive in offering their professional services to non-academic organizations. In the UK, for instance, income from consulting has increased by 50% over the past 10 years and now accounts for 10% of total external university income (HEFCE 2017). Academic consulting in this context is typically defined as an advisory service performed by academics who apply their scholarly expertise for a non-academic organization, often – but not always – for financial compensation, and without the creation of new knowledge (Perkmann and Walsh 2008; Amara et al. 2013; HEFCE 2017). Academic consulting is not a new phenomenon and has played an important role in the rise of American industry and academia (Shimshoni 1970; Lowen 1990). Today it is increasingly conspicuous amongst academics in the US and in Europe (Perkmann et al. 2013) and highly valued in industry and government as a means to gain insights into academic research (Cohen et al. 2002; Bekkers and Bodas Freitas 2008; Haucap and Thomas 2014).

Despite its importance, evidence regarding the effects that consulting has on academic research is still sparse. Private and public organizations gain prominence in academia through consulting, by providing income to academics or their institutions and by shaping or inspiring research agendas, with potential consequences for academics. Prior literature on university-industry knowledge transfer has examined the potential influence that private sector involvement can have on academic research, raising concerns for openness and the pursuit of fundamental research (Boyer and Lewis 1984; Blumenthal et al. 1996; Thursby et al. 2007) while also acknowledging positive spillovers, including ideas and revenue for research (Lee 2000; Perkmann and Walsh 2008; Buenstorf 2009).

Less studied, albeit very widespread, is academics' involvement with the public sector. We could suspect that insights from the knowledge transfer literature can also be applied to the case of public sector consulting, however, the services expected by public and private organizations can differ substantially as may financial compensations. Moreover, public and private sectors clearly differ in their management and organization with potentially different ramifications for academics working with these sectors (Boyne 2002). Recognising that these two consulting modes may have differing effects on research is crucial for defining government and university policy.

In this paper we thus focus on two forms of academic consulting, which coexist and comprise different types of knowledge and services (consulting with the public and with the private sector), and investigate the relationship between consulting and research outcomes. In order to do so, we firstly need to understand whether there are differences in the types of academics that provide advice to the public and private sectors. Again, while the drivers of private sector consulting have been discussed extensively within the context of university-industry interactions (e.g. Klofsten and Jones-Evans 2000; D'Este and Patel 2007; Link et al. 2007; Perkmann and Walsh 2009;

Jensen et al. 2010; Grimpe and Hussinger 2013), much less is known about public sector consulting. The few insights into overall public sector engagement suggest that it is more widespread in the medical and social sciences disciplines (Hughes et al. 2016), fields that have been found to engage little with the private sector, suggesting that we could expect different selection effects compared to private sector consulting.

First evidence regarding the effect of academic consulting on publication numbers, comes from the US (Rebne 1989; Mitchell and Rebne 1995) and Spain (Rentocchini et al. 2014). While the former find a positive but marginal effect for consulting time on publications at low to moderate levels, the latter reports a negative effect for very high amounts of income generated from consulting. These studies have some limitations: The US studies only consider private sector interactions and do not control for selection into consulting. The Spanish study, on the other hand, only considers university income generating forms of consulting. However, we know that academics often work with industry directly, bypassing their university (Bodas Freitas et al. 2013), and that consulting can happen pro bono and therefore does not always create a reliable paper stream (Amara et al. 2013; Perkmann et al. 2015).

Our analysis builds on data from a sample of more than 900 academics in Germany in various disciplines and makes use of survey information on academics' work time distributions in a usual workweek to identify the occurrence and intensity of different consulting activities. In terms of research outcomes, we study publication numbers and citations to publications for those who stay research active. In addition, we consider the outcome of zero publications as an exit from academic research as potential consequence from consulting. In our analysis, we observe consulting academics to have a higher probability to cease publishing altogether, but do not find lower ex-post scientific publication numbers for those who do not exit. Moreover, consulting to the public sector is associated with lower average citation numbers, which may indicate publications of less relevance for academic research. We observe disciplinary as well as academic rank differences, which we attribute to differences in career opportunities and research spillovers that can be realized.

3.2 Consulting and research outcomes

The effect of consulting to the public and private sector

Public debate repeatedly centered on the possible impact that consulting activities with public and private organizations may have on academic research outcomes, including scientific publications, research agenda setting, collaborative research or probability to exit from academia (Erk and Schmidt 2014; OECD 2015). Theoretical arguments underpinning much of the literature on university-industry interactions have generally argued that academics face time-allocation issues leading to trade-offs for research unless spillovers can be utilized (Jensen et al. 2010; Bianchini et al. 2016). Still, despite calls for more empirical evidence, little attention has been given to the investigation of consulting and its potential research spillovers.

Most empirical studies to date have considered wider knowledge transfer activities with industry, which may include consulting, contract research, academic patenting and entrepreneurship, in their investigation. This literature largely found that academic patenting and academic entrepreneurship are positively related to research performance (van Looy et al. 2006; Breschi et al. 2007; Thursby et al. 2007; Fabrizio and DiMinin 2008; Azoulay et al. 2009; Buenstorf 2009; Czarnitzki and Toole 2010). The positive spillover effect has been linked to research ideas obtained through the involvement in more applied research projects or financial benefits from commercialization that feeds positively into academic research (Lee 2000; Breschi et al. 2007; Buenstorf 2009). Sceptics, instead, have argued that engagement in knowledge exchange activities may result in late- or non-dissemination of research results (Blumenthal et al 1996; Florida and Cohen 1999; Krimsky 2003; Czarnitzki et al. 2015) or in applied research agendas that are less suitable to journal publications (Etzkowitz and Webster, 1998; Vavakova 1998; Hottenrott and Lawson 2014). Empirically, several studies looking at the effect of collaborative and contract research income on research productivity[4] find that it leads to fewer publications or fewer citations per paper, thus providing some evidence for potentially negative spillovers (Manjarrés-Henríquez et al. 2009; Hottenrott and Thorwarth 2011, Banal-Estañol et al. 2015).

Results from the few existing empirical studies that explicitly explore the influence of academic consulting on research performance suggest that, at least in the case of private sector consulting, it does not compromise academic research, at least up to a certain threshold. For example, Rebne (1989) and Mitchell and Rebne (1995), studying consulting amongst US academics, find a positive relationship at low to moderate levels of time spent on consulting with industry, but a decline in publications at high levels. More recently, in the case of academics in Spain, Manjarrés-Henrìquez et al. (2009) and Rentocchini et al. (2014) find a negative effect of consulting on publications, if a considerable amount of income is generated from it. These results suggest that consulting activities, particularly at high engagement intensities, may crowd out research activities, but also that consulting can complement publications up to a certain threshold.

The link between public sector consulting and research performance has instead not yet been explicitly explored. The nature of public sector consulting can be quite different from interactions with private firms with implications for the extent to which research spillovers can be realized. Public consulting often serves the purpose of supporting government decisions ex-ante or

[4] Sponsorship from the private sector in particular may include income from consulting projects with firms and therefore indirectly reflect an academic's engagement in consulting activities with the private sector. In addition, consulting and contract research for industry are highly correlated (Gulbrandsen and Smeby 2005).

evaluating government policies ex-post. It also often involves submitting recommendations or developing guidelines (OECD 2015). Academics may also be called on to serve on expert committees (OECD 2015), such as the Council of Economic Advisors (CEA) where economists provide direct consulting to the US government, or the Standing Committee on Immunisation (STIKO), which is composed of medical experts and provides recommendations concerning vaccination schedules in Germany. Still, the potential for cross-fertilization in terms of ideas and funding may be low in consulting activities with the public sector which is more likely reputation-based and focused on past expertise rather than addressing problems at the research frontier. Public sector consulting may therefore to a lesser extent be linked to a specific skill or current research project of an academic compared to private sector consulting, which is more about technology- or problem-specific knowledge. Translational skills are needed for both types of consulting, but in the case of private sector consulting translation may go from basic to applied research (Hottenrott 2012) while in public sector consulting academics translate research into policy or layman's terms (Salter 1988; Jasanoff 1990). Thus even though insights into policy problems may have the potential to result in scholarly articles as well as revenue for academics (Jacobson et al. 2005), the problems may be rather context-specific or of local relevance and revenues from public sector consulting may be less substantial compared to income generated with the private sector. Overall, they may thus be less effectual at supporting academics' overall research efforts through cross-funding.

Based on these arguments, we expect the trade-offs between consulting and research and the effects on publication numbers to be similar for private and public sector consulting. In terms of scientific quality or general scientific relevance, as indicated by citations to research articles, this may imply that public consulting comes at the price of fewer citations. Private sector consulting may also result in more applied research, but still be relevant to, and thus cited by, the applied research community.

At the far end, i.e. when a large share of time is dedicated to consulting, the negative spillovers may result in an exit from academic research. Specifically, in pursuing outside activities, academics may stop academic research to engage full time in other occupations including consulting, board services or spin-off creation. This exit can be due to insufficient relevance of consulting for research or time constraints that no longer allow for the pursuit of publishable research, such as a full-time move into consulting (Czarnitzki and Toole 2010; Toole and Czarnitzki 2010; Hottenrott and Lawson 2014; 2017). Hottenrott and Lawson (2017) show, for instance, that university departments that engage in contract research with industry are more likely to see departing academics move to the private sector or to non-research work within the public and university sector. Consulting may thus be conducive to a move out of academia or the take-up of more administrative or advisory posts within the university or research institute, activities that would not result in publications in academic journals.

Discipline and academic rank as moderating factors

In the discussion of research spillovers, it is important to consider that engagement in consulting does not occur at random. This becomes particularly apparent when comparing disciplinary fields or academic ranks (Bianchini et al. 2016). In engineering the share of academics engaged in private sector or paid consulting is particularly high when compared to other fields (D'Este and Patel 2007; Landry et al. 2010; Rentocchini et al. 2014). A 2015 survey of more than 18,000 UK academics, for example, found that 44% of academics in engineering provided consulting services in the previous three years, compared to just 25% in natural sciences or the humanities (Hughes et al. 2016). The same survey, however, finds that public sector engagement and advisory board services are particularly relevant for groups that have been found to engage little with the private sector, such as social and medical sciences. Consulting has moreover been linked to seniority, with the most senior academics having more opportunities to engage in consulting regardless of sector, most likely for reputation reasons (Link et al. 2007, Boardman and Ponomariov 2009; Amara et al. 2013; Rentocchini et al. 2014).

The non-random engagement in consulting has consequences for its spillovers onto research. The groups of academics that have more opportunities to provide consulting, for instance the more experienced and those in more applied fields, may be able to generate more positive spillovers from their consulting work (Bianchini et al. 2016) as they may be better able to link consulting to their research, and thus be less likely to compromise their publishing activities. This means that for these academics consulting should be less likely to lead to a reduction in the number of publications or citations, compared to those that have fewer engagement opportunities, i.e. the younger and those in more basic science field.

Again, at the far end, i.e. when a large share of time is dedicated to consulting, these groups may be more likely to exit from academic research as discussed above. The probability to exit from academic research has generally been attributed to a low "taste for science" (Roach and Sauermann 2010; Balsmeier and Pellens 2014) or the attractiveness of the private sector compared to the academic one (Stephan 2012). These attributes relate heavily to external demand and time-allocation and are likely to differ by disciplinary field and academic rank. Academics in fields that provide ample opportunity for consulting may have a lower taste for science relative to other academics and see more opportunities outside of research. They may therefore ex-ante be more likely to exit from publishable research. Moreover, private sector organizations typically pay better for highly specialized scientific expertise raising the opportunity costs of a research career (Agarwal and Ohyama 2013; Balsmeier and Pellens 2016) especially in science and engineering (BUWIN 2017, p. 182-183). Moreover, academics close to the end of their career may cash in on their experience and reputation through engaging in consulting or other less research-oriented activities at the expense of publishing (see Bianchini et al. 2016; Zucker et al. 2002). In terms of career progress there are usually no disadvantages to the decision of focusing on non-research related tasks for senior and tenured academic staff in countries such as Germany. Younger academics at the start or training phase of their career face a different effort allocation problem. While one could argue that their opportunity costs for leaving academic research are lower, they usually also have fewer opportunities to engage in consulting. However, those, that are not yet decided on a specific career or have an overall lower taste for science (Balsmeier and Pellens 2014), may find that consulting raises their employability outside academia and thus are more likely to exit from research.

To summarize, a researcher's discipline and rank may moderate the effect of consulting on research outcomes and the likelihood to cease publishing altogether. At the high end of consulting we expect junior researchers who are not yet settled on an academic career and very senior academics who have more outside opportunities to be more likely to exit from academic research, while at the low-to mid-range we expect senior academic staff to generate more positive spillovers. Further, academics in engineering may be more likely to generate positive research spillovers compared to those in the social sciences or more basic science disciplines but are also expected to be more likely to exit from academic research due to better outside opportunities and demand.

3.3 Data and model specification

Data

We build on data from a survey of academics in Germany at both, universities and non-university public research organizations (PROs).[5] The survey was conducted by the Centre for European Economic Research (ZEW) in 2008 and targeted academics in the humanities and social sciences, engineering, life science and natural sciences. Researchers were contacted by email. Contact information on university researchers was obtained from the "Hochschullehrerverzeichnis" which is a register of university personnel. Email addresses for researchers at PROs (Fraunhofer Society, Max Planck Society, Helmholtz Association, and Leibniz Association) was collected using internet search. This yielded a sample of 16,269 researchers of which 2,797 responded to the survey (including incomplete responses). Survey questions referred to the pre-survey period from 2002 to 2008 or to the current year. We complemented the survey data with publication data from Thomson Reuters Web of Science (WoS). In particular, we performed text field searches on the academics' names in the publication database (articles, books, reviews, proceedings) and manually screened matches based on CV and website information. Further, we searched the

[5] PROs play an important role in the German academic research landscape. PROs include the Fraunhofer and Max-Planck Society, as well as the Helmholtz- and Leibniz Associations, and accounted for around 20% of academic staff in 2012 and for 34.4% of the European Research Council grants awarded to German institutions during the period 2007-2013 (DFG 2015).

Espace database of the European Patent Office (EPO) and the database of German Patent Office for patents (DPMA) on which the academics appear as inventors. As in the case of publications, all matches were manually checked. Eventually, we obtain publication and patent records for all individual academics from 2002 until 2013 and citations to their publications until autumn 2015. In our cross-section of academics, publications are collected for a pre- and post-survey period. The collection window, and thus the citation time windows, are identical for all surveyed academics. The censoring of citations to newer articles should thus be of minor concern. - Removing observations with incomplete records in the survey questions, the final sample comprises 951 individual-level observations. Table 3.1 reports descriptive statistics for all the variables used in the analysis (for pairwise correlations see Table 3.6 in the Appendix).

Representativeness of the sample

To check for the representativeness of our sample we compare it to the German academic population as a whole in terms of institution type, gender, discipline and age (see Table 3.7 in the Appendix). Aggregate information on the academic population was collected from the Federal Statistical Office data base (DESTATIS). The sample distribution differs somewhat from the population in terms of institution types because of an overrepresentation of PROs, an intentional aspect of the survey frame. In terms of disciplinary fields there are only small differences between the sample and the population. In terms of age classes, we find that younger researchers are underrepresented in the data, which may also contribute to the overrepresentation of males. The underrepresentation of young researchers partially stems from the fact that surveyed academics were identified using a list of university staff, the *Hochschullehrerregister*, which only lists few junior academics. The differences observed between the population and the survey respondents in terms of institution type, age and gender are therefore assumed to not represent a non-response bias. Still, to address these sample characteristics, we construct field-institution type weights to capture some of the observed differences (see also Czarnitzki et al. 2015). We apply inverse probability weighting using population weights to test the robustness of our results to these sample properties. Comparing these to the results of the unweighted models, we observe some small differences in the estimated coefficients, but these differences do not qualitatively change the results (compare Tables 3.3 and 3.4 to Tables 3.9 and 3.11 in the Appendix).

Dependent variables

The main variables of interest are the research performance of academics in the post-survey period (2009 to end 2013) and their (temporary) exit from publishing. We consider the exit from research work to be reflected in zero WoS publications in the five year post-survey period (*exit*). This variable thus reflects publication inactivity over that period and not necessarily the termination of a work contract. About 18% of academics have no publications in WoS in the post-survey period, while the average number of *publications* is 12.4 and each publication receives about 12 citations (*average citations*) in the time window considered. From the individual publication and citation counts, we further derive field-weighted counts to account for heterogeneous publication/citation patterns of different disciplines. To obtain these values we divide publication counts as well as average citations by the within-sample field averages (*field-weighted publications, field-weighted average citations*). A value below one represents a below field-average output and a value above one represents an above field-average output.

Consulting activities

Our data is distinctive from previous studies in using the time share that academics devote to consulting (*consulting*). The advantage of using survey-based time shares as opposed to consulting income or official university records[6] is that academics have no incentives to under or over report their consultancy work. In addition, we capture consulting activities for which no financial compensation had been received. Despite the downsides in terms standardization and recall difficulty in surveys, we avoid problems in measuring consulting activities that arise if individuals are able to charge very different fees and thus have different levels of income per hour of consulting work. It also captures activities that do not leave a paper trail. The consulting time-share refers to a typical work week and is therefore cross-sectional in nature. Based on the survey responses, we distinguish between consulting to the private (*private consulting*) and the public sector (*public consulting*).[7] This is different from Rebne (1989) and Mitchell and Rebne (1995) who use the number of hours spend on consulting work or Rentocchini et al, (2014) who rely on consulting income.

Table 3.1 shows that academics spend roughly 5.3% of their time on consulting, on average. Among consulting-active academics the average time spent on consulting is 12.2%. By comparison, about 50% of time is spent on research, and 21% on each teaching and administration (see Table 3.8 in the Appendix for more details on time distributions).[8] While the overall time-share devoted to academic consulting is not high, 44% of academics reportedly engaged in some form of consulting at the time of the survey; about 17% provide consulting only to the public

---

[6] While German law in principle requires research staff at universities and PROs to report additional consulting income to their employer, there are certain exemption levels that vary between different institutions below which no reporting is required (Hochschul-Nebentätigkeitsverordnung, HNtV). Thus, income information provided by institutions would not provide a full picture.

[7] The questionnaire asked: "Please give the percentage of working time you currently spend on the following activities." Respondents distributed timeshares over: research, research funded by research grants, teaching, administration, private sector consulting and public sector consulting. Unlike research funded by research grants the general research category refers to research financed by institutional core funding which, in the case of Germany, is typically distributed to the universities or PROs through the state and is not subject to a specific project proposal, application or selection process. See Table 3.8 in the Appendix for an overview of the time share distributions.

[8] By comparison, a 2015 survey of academic staff in the UK found that academics spend about 40% on research, 30% on teaching and 21% on administrative tasks (Hughes et al. 2016). The lower teaching share in our survey will be primarily due to the additional surveying of PROs rather than country differences.

sector, 13% only to the private sector, and 14% to both. A detailed comparison of consulting active and inactive academics is provided in section 3.2.

Moderators

Of the academics in the sample, 21% belong to *social sciences* (and humanities), 30% to *life sciences* (biology, medicine, agriculture and veterinary sciences), 31% to the *natural sciences* (chemistry, physics, earth science and mathematics) and 19% are active in *engineering*. More than half of the sample are employed as *professors* (54%), 11% are *assistant professors* (including academics working towards habilitation), 26% are *senior researchers* and about 10% are *junior researchers* (scientific assistance staff that do not hold and/or are studying for a PhD).

Controls

A series of other controls are included that have been shown to affect publication outcomes, such as age and gender (e.g. Toole and Czarnitzki 2010; Mairesse and Pezzoni 2015). Academics are, on average, 49 years old (*age*), and 15% are *female*. More than half of the academics in the sample (59%) are employed at universities (*university*), while the rest work at PROs or other research institutions. We also include variables that capture the effect of network and funding on academic output. This includes the size of the local peer group in terms of the number of people from the same institution working in closely related fields (*peer group size*), a measure for *collaborative reach* based on the location of research partners during the 2002 to 2008 period[9], and a measure for *international visibility* based on reported international conference participation during an average year. The survey also includes information on academics' grant-based research income from the European Union, national and regional governments, science foundations, such as the German Research foundations (DFG), industry and other external funders during the period 2002 to 2006. Funding amounts are aggregated into, *industry funding* and *public funding*. Finally, we include a binary indicator for co-authored articles with employees from the private sector in the previous 12 months *(co-authorship industry)*, and the number of patents in the pre-survey period (*patents)* as additional controls.

All regression models also include pre-survey publication and citation numbers (between 2002 and 2008) as predictors of future publication performance. In addition, we control for the average number of co-authors on publications in the pre-survey period (*average number of co-authors*). Academics published on average 12 items in the pre-survey period and received an average of 24 citations per publication. The average number of co-authors is four with the lower values in social sciences (1.2) and engineering (3.4) compared to life sciences (5.4) and natural sciences (7.8).

---

[9] The variable takes values from zero to five, where zero stands for "no collaborative work", one for "collaboration only within the home institution", two for "collaboration only inside Germany", and three for "European-wide collaboration, but not beyond". Categories four and five capture collaboration with North America and the rest of the world, respectively.

Table 3.1: Descriptive statistics

Variable	unit	source	median	mean	s.d.	min.	max.
Outcome Variables							
$\text{exit}_{2009\text{-}2013}$	count	WoS	0	0.18	0.38	0	1
$\text{publications}_{2009\text{-}2013}$	count	WoS	6	12.44	20.13	0	278
av. $\text{citations}_{2009\text{-}2013}$	fraction	WoS	8.44	11.85	15.82	0	157.67
field-weighted $\text{publications}_{2009\text{-}2013}$	fraction	WoS	0.53	1	1.57	0	16.93
field-weighted av. $\text{citations}_{2009\text{-}2013}$	fraction	WoS	0.68	1	1.57	0	23.14
Consulting activities							
consulting [yes / no]	binary	Survey	0	0.44	0.50	0	1
public consulting [yes / no]	binary	Survey	0	0.31	0.46	0	1
private consulting [yes / no]	binary	Survey	0	0.27	0.44	0	1
consulting	percentage	Survey	0	5.31	10.27	0	100
public consulting	percentage	Survey	0	3.06	7.96	0	100
private consulting	percentage	Survey	0	2.25	6.23	0	100
Moderators							
junior researcher	binary	Survey	0	0.09	0.29	0	1
senior researcher	binary	Survey	0	0.26	0.44	0	1
assistant professor	binary	Survey	0	0.11	0.32	0	1
full professor	binary	Survey	1	0.54	0.50	0	1
social sciences	binary	Survey	0	0.21	0.41	0	1
life sciences	binary	Survey	0	0.30	0.46	0	1
natural sciences	binary	Survey	0	0.31	0.46	0	1
engineering	binary	Survey	0	0.19	0.39	0	1
Controls							
age	count	Survey	49	49.40	8.28	28	74
female	binary	Survey	0	0.15	0.36	0	1
$\text{publications}_{2002\text{-}2008}$	count	WoS	4	11.70	21.03	0	305
average $\text{citations}_{2002\text{-}2008}$	fraction	WoS	16.13	24.18	31.67	0	344.2
field-weighted $\text{publications}_{2002\text{-}2008}$	fraction	WoS	0.47	1	1.81	0	24.52
field-weighted average $\text{citations}_{2002\text{-}}$	fraction	WoS	0.69	1	1.39	0	17.18
average number of $\text{co-authors}_{2002\text{-}2008}$	fraction	WoS	4.46	5.99	17.30	0	332.83
collaborative $\text{reach}_{2002\text{-}2008}$	ordinal	Survey	3	3.06	1.36	0	5
international visibility	fraction	Survey	0.71	0.69	0.17	0	1
industry $\text{funding}_{2002\text{-}2006}$	amount	Survey	0	0.16	0.46	0	11
public $\text{funding}_{2002\text{-}2006}$	amount	Survey	0.40	1.10	3.03	0	75
peer group size	count	Survey	10	39.46	148.47	0	3,000
university	binary	Survey	1	0.59	0.49	0	1
$\text{patents}_{\text{pre2009}}$	count	EPO/DPM	0	1.06	3.72	0	41
co-authorship industry	binary	Survey	0	0.22	0.41	0	1
Exclusion restrictions							
regio skills	percentage	INKAR	9	10.01	6.03	0.70	43.80
firm	binary	Survey	0	0.17	0.38	0	1
techtransfer industry	binary	Survey	0	0.43	0.50	0	1

Notes: Number of observations is 951. Funding variables in 100.000€. There are two individuals with consulting shares of 100%, one for each type of consulting. Both are project leaders so that the answer seems realistic and not a measurement error. The reference period for the citation variables (for instance 2009-2013 or 2002-2008) refers to publications in that period and the citations received by these publications up to autumn 2015.

Descriptive analysis of consulting activity

Table 3.2 compares the mean values of the dependent variables (publications, citations and exit), and the moderators academic rank and discipline by consulting activity. We observe a higher average number of publications, but fewer citations for consulting active researchers and no significantly different share of "exits". In addition to the mean comparisons, Figure 3.1 shows the number of publications and average citations per publication (in the post-survey period) over different percentiles of the consulting time-share distribution. For both variables and both types of consulting, research output, in particular the median, tends to be lower at higher time-shares spent on consulting. These descriptive statistics suggest that not consulting engagement as such matters, but the intensity of the engagement.

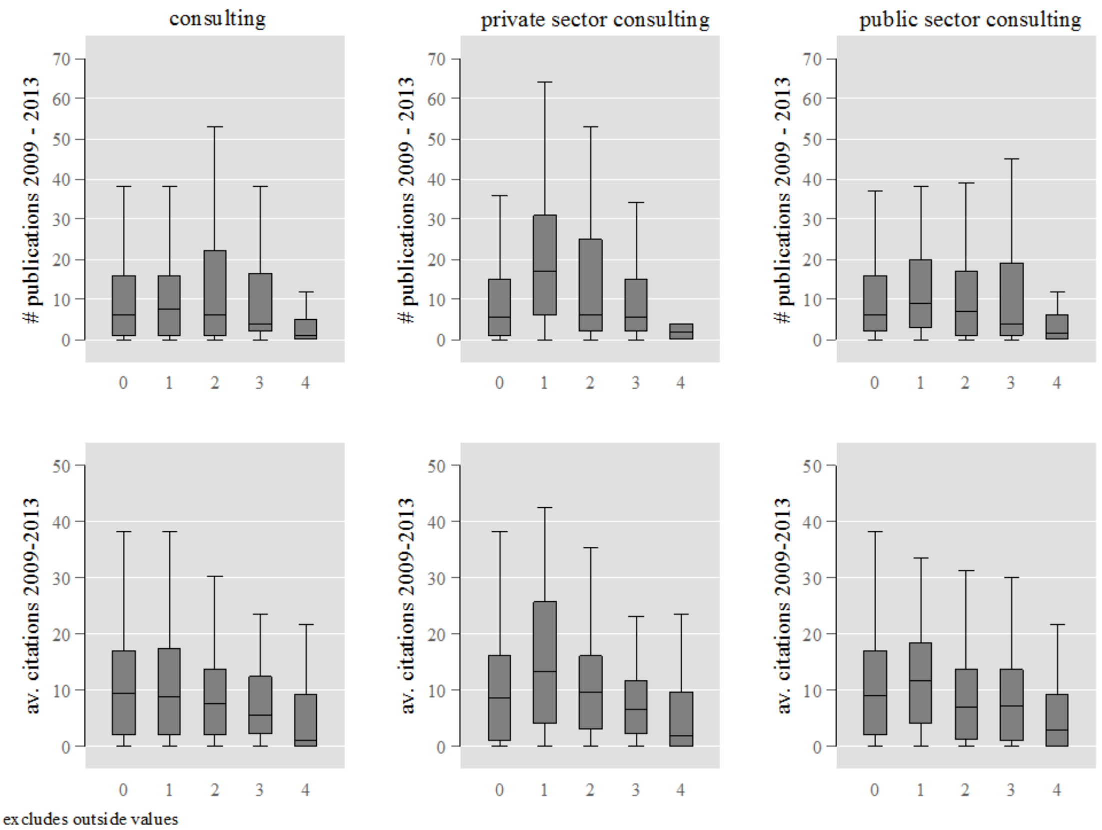
Table 3.2: Descriptive statistics by type of consulting (selected variables)

	No consulting I.	Consulting active II.	Private sector consulting III.	Public sector consulting IV.	I. vs. II.	I. vs. III.	I. vs. IV.
Observations	537	414	255	292			
	mean (s.d.)	mean (s.d.)	mean (s.d.)	mean (s.d.)		t-test	
Outcome variables							
exit$_{2009\text{-}2013}$	0.17 (0.37)	0.19 (0.39)	0.13 (0.34)	0.19 (0.39)	0.46	0.91	0.38
publications$_{2009\text{-}2013}$	11.33 (17.63)	13.88 (22.90)	15.55 (24.24)	13.58 (22.25)	0.05	0.01	0.11
av. citations$_{2009\text{-}2013}$	12.72 (16.99)	10.71 (14.11)	11.41 (14.85)	9.93 (11.84)	0.05	0.29	0.01
field-weighted publications$_{2009\text{-}2013}$	0.88 (1.44)	1.15 (1.73)	1.26 (1.73)	1.14 (1.74)	0.01	0.00	0.03
field-weighted av. citations$_{2009\text{-}2013}$	1.01 (1.51)	0.99 (1.65)	1.02 (1.68)	0.92 (1.36)	0.90	0.92	0.42
Moderators							
junior researcher	0.10 (0.31)	0.08 (0.27)	0.06 (0.24)	0.09 (0.29)	0.24	0.06	0.48
senior researcher	0.26 (0.44)	0.25 (0.43)	0.27 (0.44)	0.21 (0.02)	0.67	0.86	0.12
assistant professor	0.15 (0.36)	0.06 (0.24)	0.05 (0.01)	0.06 (0.01)	0.00	0.00	0.00
full professor	0.48 (0.50)	0.61 (0.49)	0.62 (0.49)	0.64 (0.48)	0.00	0.00	0.00
social sciences	0.19 (0.40)	0.23 (0.42)	0.14 (0.35)	0.26 (0.44)	0.18	0.07	0.02
life sciences	0.28 (0.45)	0.32 (0.47)	0.31 (0.46)	0.33 (0.47)	0.29	0.54	0.16
natural sciences	0.38 (0.49)	0.21 (0.41)	0.21 (0.41)	0.22 (0.41)	0.00	0.00	0.00
engineering	0.14 (0.35)	0.24 (0.43)	0.34 (0.48)	0.18 (0.39)	0.00	0.00	0.09

Notes: 133 researchers (14%) engage in both public and private sector consulting. Two-sided t-tests presented [Pr(|T| > |t|)].

Table 3.2 further shows that in the social sciences, public sector consulting is more prevalent than private sector consulting or no consulting, while in life and natural sciences the differences are less pronounced. In natural sciences, we observe the overall lowest involvement in consulting. In engineering, consulting with the private sector is reported by about 34% of academics (see Table 3.8 in the Appendix for consulting time-shares and the share of consulting-active academics by discipline). Looking at academic rank, we see that the share of full professors is largest in all consulting groups and also significantly larger than in the non-consulting active sub-sample. Also a large share of senior researchers is engaged in consulting, with little differences between types, while assistant professors are least represented in all consulting types. For junior researchers public consulting is slightly less common than private sector consulting or no consulting.

Figure 3.1: Box plots of outcome variables over consulting time-shares (951 observations)



Notes: Percentiles defined as ranges based on the consulting time-share percentiles for consulting-active researchers 1 to 10 = 1, 11 to 50 = 2, 51 to 90 = 3 and > 90 = 4. Graph colour scheme from Bischof (2016).

It is moreover interesting to point out, that certain attributes differ considerably between consulting-active and non-consulting active academics. While the former spend significantly less time on block-funded research (17% versus 23%) and less time on grant-based research (30% versus 34%), teaching loads differ only slightly (20% versus 23%) and administrative duties are similar (both 21%). These numbers suggest that consulting may substitute research, but is not associated with a higher administrative burden or less time devoted to teaching (see Table 3.8 in the Appendix).

Estimation Strategy

We estimate the probability of exit and publication performance while accounting for selection into consulting. Engagement in consulting does not occur at random and modelling the selection into consulting enables us to correct for the selection bias in consulting activity. Moreover, we prefer selection type models over other treatment effects models as they allow to follow a suggestion by Wooldridge (2002, p. 594) to include the logarithm of an academic's pre-sample research performance[10] in the outcome equations to capture i) path dependency and cumulative advantage effects in publication and citation numbers and ii) the otherwise unobserved ability to publish of an individual academic. These initial performance variables proxy for permanent individual unobservable effects, or "fixed" effects, which are not directly observable, but associated with underlying variables, including individual capability, motivation and talent (Mairesse and Pezzoni 2015). Finally, our modelling approach also has the advantage that we can explicitly model the propensity to engage in consulting. The results from the selection stage are informative as such and also enable a closer comparison to the existing academic consulting literature.

In the model's selection equation we include personal and institutional attributes which have been shown to be of relevance in explaining academic consulting in several previous studies. In addition, the selection equation includes a set of exclusion restrictions which help to identify the second stage. These are the share of employment in knowledge-intensive industries in a region (*regio skills*), and commercial activities that have been linked to consulting such as firm foundation experience (*firm*) and technology transfer activities with industry (*techtransfer industry*) during the previous 12 months (the latter two are based on survey responses). The share of employment in knowledge-intensive industries is calculated at the 4-digit municipality level based on data from the INKAR database provided by the German Statistical Offices in cooperation with The Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). The skill-wise labor market composition in a region may determine the demand for academic consulting services both for the public and private sector, but not affect an individual researcher's publication performance. Founder experience (unlike current entrepreneurial activity) may reflect networks that facilitate consulting, but does not directly correlate with the output measures. Likewise, technology transfer through means other than consulting create networks to the private sector and generate consulting opportunities, but does not necessarily affect publications in a particular direction. We test the statistical appropriateness of these exclusion restrictions in auxiliary regressions which show that the excluded variables are individually and jointly insignificant in the outcome equations, but indeed relevant in the selection equation.

[10] The pre-sample variables are adjusted to the respective dependent variable, i.e. based on field-weighted publication counts if the dependent variable is ln(field-weighted publications+1) and field-weighted average citations in the model for ln(field-weighted av. citations+1).

The selection into consulting is thus estimated for each academic *i* as:

$$\Pr(consulting)_i = \beta_0 + \sum_{n=1}^{3} \beta_n er_i + \sum_{n=4}^{k} \beta_n controls_i + u_i \qquad (1)$$

with the vector *er* referring to the set of exclusion restrictions, *k* is the total number of regressors and parameter u is the error term.

We then proceed in two steps, differentiating between the effects of consulting on exit and on research performance. In the research performance models, we exclude individuals with zero publications in the five-year post-survey period since we consider these as no longer research active. Their zero publication output is captured in the exit models and including them in the research performance equations would confound reduced output of research-active academics with those that are research in-active. It should be noted here, that inactivity in terms of publications is defined over the relatively long period of 5 years and thus does not apply to someone with a publication break of just a single year (or two, or three). If consulting indeed leads to a higher probability of exit, we would potentially overestimate the (negative) effect of consulting on research output of those who remain research active due to the zero publication counts.

We thus resume in two steps as follows. We firstly estimate the probability of exit from academic research while accounting for the selection into consulting (as specified in equ. 1) using a Heckman-type procedure for binary outcome variables estimated by maximum likelihood method (van de Ven and van Pragg 1981; De Luca 2008). Exit probability is then modelled as follows:

$$\Pr(exit)_i = \gamma_0 + \gamma_1\, consulting\ share_i + \gamma_2\, consulting\ share_i^2 + \sum_{n=3}^{k} \gamma_n\, controls_i + \alpha^{\rho} + \tilde{u}_i \qquad (2)$$

with $corr\,(u, \tilde{u}) = \rho_{u,\tilde{u}}$. A statistically significant $\alpha^{\rho} = 0.5 \ln(1+\rho)\,/(1-\rho)$ indicates that some selection bias would be ignored in the absence of the selection equation. In addition, we estimate models with interaction effects between the consulting share and the moderators academic rank and disciplinary field.

This model is first estimated for overall consulting, before we specify a model in which we explicitly distinguish time devoted to public sector versus private sector consulting. The second order term is included to account for possible non-linear effects. The vector *controls* includes the academics' age, a gender dummy, a university dummy, field-weighted publications and field-weighted average citations in the pre-sample period, patents, grant-based research funding, scientific field and rank dummies.

Next, we estimate research performance in terms of publications and citations for those academics that remain research active using linear endogenous switching models (LES). LES models are a variant of the selection model (see Lokshin and Sajaia 2004) that account for the non-randomness of consulting activity in the effect of consulting on post-survey research performance. Unlike Heckman-type correction models, LES models estimate the outcome equation for both groups of the selection. This means they also provide an outcome equation for consulting *in*active academics, allowing for a comparison of control variables between the two groups.

As above, we estimate separate models for consulting in general and the two types of consulting, and for the different publication-based outcome variables:

$$\ln(outcome)_i \; = \; \gamma_0 + \gamma_1 \, consulting\ share_i \; + \gamma_2 \, consulting\ share_i^2 + \sum_{n=3}^{k} \gamma_n \, controls_i + \alpha^{\rho} + \tilde{u}_i \qquad (3)$$

The consulting equation is specified according to equation (1) and is estimated jointly with the outcome equation (3) via full information maximum likelihood method (FIML) and α^{ρ} is calculated as described above. We employ the natural logarithm of the publication count and average citation numbers. Log transforming variables with skewed distributions has several advantages and is quite common in the context of publication measures (see for instance, Fabrizio and Di Minin 2008; Buenstorf 2009; Banal-Estañol et al. 2015). First, it reduces the skewness of the distribution as well as heteroscedasticity because it suppresses variation in the data and makes the error distribution more normal. Second, it makes interpretation straight forward. A one percentage point change in our consulting share can be interpreted in terms of percentage change in the outcome variable.[11] In addition to these baseline models, we again estimate models with interaction effects between the consulting share and the moderator's academic rank and disciplinary field.

In addition to the selection models that rely on the set of exclusion restrictions, we test the robustness of the results to using an instrumental variable (IV) approach suggested by Lewbel (2012). This method does not rely on external IVs, but achieves identification through the generation of IVs based on heteroscedasticity.

[11] We checked the sensitivity of the results of the publication outcome models to different estimation methods (OLS, Tobit, Poisson and negative binomial estimation) and to different specifications of the dependent variable (levels versus log transformation). These tests showed that estimated coefficients are quantitatively and qualitatively similar (see Table 3.12 in the Appendix).

3.4 Results

Selection into consulting

Table 3.3 shows the results (marginal effects) from the set of probit models that represent the selection equation, i.e. the probability of engaging in any consulting (model 1), and results from simultaneous probit models on public consulting and/or private sector consulting (model 2).[12] As expected, we find that academics in the social sciences are more active in public consulting than academics in science and engineering. There are however fewer differences with regard to involvement with the private sector. We further find that professors and junior researchers are more likely to be active in consulting, especially in public consulting, than mid-career researchers. Professors are also most active in consulting to the private sector. Similar findings were reported in Amara et al. (2013) who show that research staff and full professors are more likely to engage in paid consulting than mid-career academics. We further find a positive effect of age, which supports prior findings on the higher likelihood of older academic staff to engage in industry consulting (Louis et al. 1989; Boardman and Ponomariov 2009). Interestingly, the effect of age is higher for consulting with the public sector compared to the private sector. We also find that women are less likely than men to engage in private sector consulting, while there is no difference for public sector consulting. This confirms Abreu and Grinevich (2013), who find that women engage less with the private sector but more with the public sector compared to men, and is also in line with prior research on industry consulting that consistently showed lower activity for women (e.g. Link et al. 2007; Grimpe and Fier 2010).

[12] See Table 3.11 in the Appendix for corresponding estimations using population weights. More precisely, we employ field-institution type inverse probability weights that should capture some of the observed differences also in terms of gender and age, and apply inverse probability weighting to test the robustness of our results to these sample properties, especially bias caused by field or institute sampling, through population weights. The differences in estimated coefficients are minor and not qualitative in nature.

Table 3.3: Results of probit and simultaneous probit models on private and public sector consulting

Model	**1**		**2**			
Dependent variable	**consulting** [yes / no]		**public consulting** [yes / no]		**private consulting** [yes / no]	
	dy/dx	s.e.	dy/dx	s.e.	dy/dx	s.e.
Moderators						
junior researcher			*Reference Category*			
senior researcher	-0.022 ***	0.006	-0.064 ***	0.009	0.018 ***	0.003
assistant professor	-0.129 ***	0.022	-0.143 ***	0.020	-0.063 ***	0.020
full professor	0.062 *	0.036	-0.017	0.047	0.066 ***	0.021
social sciences			*Reference Category*			
life sciences	-0.097 *	0.057	-0.100 *	0.057	0.023 *	0.013
natural sciences	-0.246 ***	0.025	-0.217 ***	0.051	-0.056	0.036
engineering	-0.122 ***	0.039	-0.138 ***	0.034	0.047	0.046
Controls						
age	0.003 ***	0.001	0.005 ***	0.001	0.002 **	0.001
female	-0.005	0.022	0.035	0.029	-0.052 **	0.021
field-weighted publications$_{2002\text{-}2008}$	0.006 **	0.003	0.006	0.004	0.009 ***	0.003
field-weighted av. citations$_{2002\text{-}2008}$	-0.026 ***	0.006	-0.026 ***	0.005	-0.009	0.009
collaborative reach$_{2002\text{-}2008}$	0.016 *	0.009	0.034 ***	0.012	-0.002	0.008
international visibility	0.035	0.046	-0.003	0.040	0.081	0.054
ln(industry funding)$_{2002\text{-}2006}$	0.225 **	0.112	-0.085	0.053	0.225 ***	0.088
ln(public funding)$_{2002\text{-}2006}$	0.083 *	0.043	0.200 ***	0.015	-0.059 ***	0.019
ln(peergroup size)	-0.021 ***	0.008	-0.019 ***	0.007	-0.008	0.010
university	-0.116 **	0.053	-0.078 **	0.037	-0.042 *	0.023
ln(patents$_{pre2009}$)	-0.012	0.026	-0.061 ***	0.011	0.012	0.012
coauthorship industry	0.077 *	0.042	0.057	0.048	0.061 *	0.034
Exclusion restrictions						
regio skills	-0.006 **	0.003	-0.008 **	0.003	-0.005 *	0.002
firm	0.042 **	0.018	0.050	0.031	0.052	0.034
techtransfer industry	0.212 ***	0.020	0.073 ***	0.016	0.255 ***	0.011
Log pseudolikelihood	-560.226		-921.189			
ρ[equ. 1/2]	-		0.522 (0.046)***			

Notes: Number of observations = 951. Average marginal effects presented. *** (**,*) indicate a significance level of 1% (5%, 10%). All models contain a constant. If we include unweighted publication and citation variables the signs and significance levels are similar.

In terms of pre-survey scientific activity, we see that field-weighted average citations are negatively correlated with consulting, whereas publication counts show a positive correlation. Industry funding correlates strongly and positively with private sector consulting and negatively with public sector consulting. The contrary is the case for public funding which contradicts previous research by Jensen et al. (2010) and Muscio et al. (2013) who stressed that public funding can be a facilitator for research contracts and consulting with the private sector (see also D’Este et al. 2013; Amara et al. 2013). The findings further show that collaborative reach correlates positively with public sector consulting. The local peer group size is negatively associated with public consulting, suggesting that academics working in isolated areas are more likely to look for external consulting options. Patenting academics are also less likely to engage in public consulting. Finally, co-authorship with industry correlates positively with private sector consulting, confirming prior findings in the field (Louis et al. 1989; Landry et al. 2010). The correlation between the public and private sector consulting equation is positive and significant, pointing to the importance of estimating these equations jointly.[13] It also indicates that academics make use of both engagement modes simultaneously.

Consulting and the probability to exit from academic work

Table 3.4 presents estimated coefficients from the models on exit from academic research. In models 3 and 4 we account for the possibility of retirement and check the sensitivity of the results to the exclusion of individuals who were 64 years or older at the time of the survey. In line with our expectation, we find that consulting increases the propensity to exit from publishing. This is in keeping with studies that report exit following other forms of non-research activities such as academic entrepreneurship (Toole and Czarnitzki 2010). The average marginal effect (AME) for consulting (which cannot be directly seen from the coefficient) is 0.002 in model 1 and 0.003 in model 3 indicating that, on average, an increase in consulting by 10 percentage points increases exit probability by about 3% in the subsample of academics below the age of 65 (model 3). The effect of consulting, however, is unlikely to be the same for all consulting time-shares. Figure 3.2 therefore depicts graphically the predictive margins of consulting on exit probability at different consulting shares. We find that the probability to exit increases as consulting increases, but with diminishing marginal effects. The slope of the curve is determined by the marginal effects at representative values (MERs), i.e. the marginal impact of consulting on exit probability at different values of the consulting distribution, and is steepest at consulting time shares between 10 and about 20%.

Looking at private and public sector separately we find that in model 4, the AME for private sector consulting is 0.002 (significant at 1% level), while the AME of 0.001 for public sector consulting is insignificant. The graphs in Figure 3.2 show that for public sector consulting the impact on exit is initially small, explaining the smaller and insignificant AME. At larger intensities, particularly between 20 and 50 percent of time spent, exit probability increases substantially. This high exit propensity for public sector consulting may be due to academics taking on the role of brokers or full time advisors, no longer concerned with their scientific research (Haucap and Moedl 2013). For private sector consulting exit probability increases with consulting time-shares, but with decreasing marginal impact at very high intensities around the 95th percentile and beyond. For relatively common levels of private sector consulting, say 5% of time, an increase of consulting by 10 percentage points to 15% will increase exit probability from

[13] We also estimate simultaneous equation models on the timeshares devoted to public and private sector consulting. The effects of the explanatory variables are very similar to the ones in the probit models and the correlation coefficient between the timeshare equations is insignificant (see Table 3.10 in the Appendix).

8.9 to 11.1%. For an increase from 20 to 30% the marginal effect is still positive, but slightly smaller with an increase from 12.1 to 13.5% exit probability.

The average effects are thus rather small, but Figure 3.3 shows that results differ substantially by moderators (detailed regression results available upon request). Exit propensity as such is highest for junior (pre-PhD) research staff and lowest for assistant professors. Initially, exit propensity increases with consulting time-shares for all ranks, except for junior researchers (top left of Figure 3.3). At higher consulting shares, however, exit probability increases particularly for junior researchers and full professors, i.e. the groups that are also more likely to engage in consulting. For example, at a consulting time-share of 40% (90^{th} percentile), junior researchers have an exit propensity of 80%. These results show that consulting may distract junior academics from research and thus steer them away from a research career, in line with concerns voiced by the knowledge exchange literature (Blumenthal et al. 1996; Florida and Cohen 1999). The effect for junior researchers is driven by public sector consulting (see bottom left of Figure 3.3), though overall there is little increase in exit probability at lower time-shares, explaining the insignificant AME for public consulting. In the case of private sector consulting (see bottom right of Figure 3.3), curves also show minor differences for low values, with the steepest slope for senior researchers. For full professors the curve is flat up to a 20% consulting share, but positive at higher consulting shares. For instance, for an increase in the time-share from 40 to 50% the effect on exit probability increases by about 10 percentage points.

Effects also differ by disciplinary field. In the social sciences, an increase in the consulting share is associated with a higher exit probability with an AME of 0.035 (top right of Figure 3.3). In engineering the slope is flatter, but also positive for the entire range of consulting time-shares (AME = 0.022). In the publication intensive fields of life and natural science an increase in consulting increases exit probability at similar rates as in engineering (with AMEs of 0.021 and 0.024), but with constant or declining impact and a lower exit probability in absolute terms. The higher exit propensity for social sciences is contrary to what we would expect based on prior evidence (e.g. Rentocchini et al. 2014) and is most likely due to the consideration of public sector consulting and its high prevalence in the social sciences in this study.

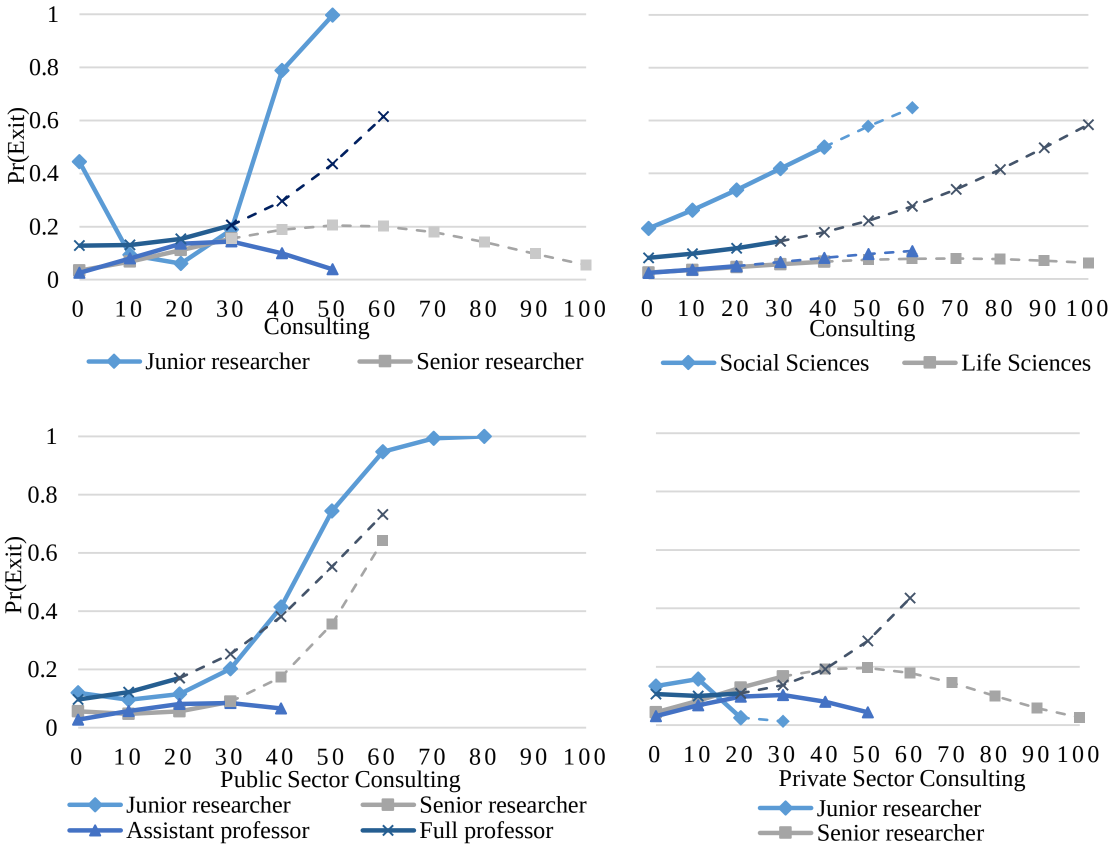
In terms of control variables, we find that exit probability increases with age. We do not find women to have a higher propensity to cease publishing, even though prior literature has attributed exit to gender and family situation (Ginther and Kahn 2004; Mairesse and Pezzoni 2015). We further find that the better the ex-ante publication performance and international visibility, as measured by conference attendance, the less likely an academic is to stop publishing. The propensity of exit from academic research also decreases with other measures of research activity, such as peer group size within the institution and patenting (Table 3.4).

Table 3.4: Estimation results from probit models with selection on "exit"

	Model 1 (full sample)		Model 2 (full sample)		Model 3 (age < 65)		Model 4 (age < 65)	
	coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.
Consulting activities								
consulting	0.025 **	0.010			0.030 ***	0.011		
consulting²	<-0.001	<0.001			<-0.001	<0.001		
public consulting			-0.003	0.014			0.007	0.018
public consulting²			0.001 *	<0.001			<0.001	<0.001
private consulting			0.024 ***	0.006			0.026 ***	0.004
private consulting²			<-0.001 **	<0.001			<-0.001 **	<0.001
Controls								
age	0.023 ***	0.005	0.026 ***	0.008	0.028 ***	0.010	0.031 **	0.012
female	-0.017	0.190	-0.031	0.210	0.007	0.187	-0.008	0.205
junior researcher				*Reference Category*				
senior researcher	-0.304 ***	0.049	-0.224 ***	0.038	-0.372 ***	0.062	-0.287 ***	0.047
assistant professor	-0.514 ***	0.128	-0.452 ***	0.108	-0.586 ***	0.138	-0.508 ***	0.132
full professor	-0.230 **	0.103	-0.252 *	0.148	-0.256 *	0.142	-0.259	0.186
field-weighted publications$_{2002\text{-}2008}$	-0.648 ***	0.038	-0.670 ***	0.060	-0.631 ***	0.036	-0.655 ***	0.059
field-weighted av. citations$_{2002\text{-}2008}$	-0.101	0.119	-0.117	0.125	-0.108	0.122	-0.121	0.133
ln(industry funding)$_{2002\text{-}2006}$	1.143	0.737	1.233 *	0.748	1.458 **	0.663	1.620 **	0.650
ln(public funding)$_{2002\text{-}2006}$	0.013	0.151	-0.035	0.164	0.074	0.115	0.012	0.118
collaborative reach$_{2002\text{-}2008}$	<-0.001	0.059	0.024	0.054	0.009	0.063	0.031	0.055
international visibility	-1.509 ***	0.270	-1.710 ***	0.263	-1.281 ***	0.189	-1.452 ***	0.273
ln(peergroup size)	-0.070 **	0.031	-0.090 ***	0.013	-0.089 *	0.046	-0.108 ***	0.030
university	0.130	0.142	0.173	0.182	0.037	0.209	0.067	0.243
social sciences				*Reference Category*				
life sciences	-0.974 ***	0.083	-1.021 ***	0.083	-1.078 ***	0.111	-1.121 ***	0.102
natural sciences	-1.106 ***	0.135	-1.171 ***	0.084	-1.090 ***	0.095	-1.145 ***	0.063
engineering	-0.570 **	0.291	-0.601 **	0.241	-0.599 **	0.288	-0.618 **	0.273
ln(patents$_{pre2009}$)	-.108 **	0.047	-0.127 ***	0.036	-0.181	0.160	-0.202	0.152
coauthorship industry	-0.205	0.224	-0.281	0.264	-0.199	0.225	-0.260	0.242
# observations	951		951		909		909	
# consulting-active obs. (2nd stage)	414		414		392		392	
Log pseudolikelihood	-686.24		-683.62		-653.92		-651.66	
Wald test of indep. equations chi^2(1)	3.09*		2.97*		4.62**		3.03*	
α^{ρ}	0.966 (0.549)*		0.731 (0.424)*		1.128 (0.524)**		0.896 (0.515)*	

Notes: Number of observations is= 951. Marginal effects at means. *** (**,*) indicate a significance level of 1% (5%, 10%). All models contain a constant. Clustered standard errors in parenthesis. If we include unweighted publication and citation variables the signs and significance levels are similar.

Figure 3.2: Predictive margins for “exit” by rank and field



Notes: Predictive margins are only shown for the range of consulting values where the margins are significant at least at 10% confidence level. Predicted exit probability depicted in the y-axis and consulting share on the x-axis. Predictive margins are calculated at the deciles of the consulting time-share distribution. The slope of the predictive margins curve represents the marginal change in the predicted probability for a change in the consulting time-share, i.e. the marginal effect. Dashed lines indicate values beyond the 95th percentile of the consulting share distribution.

Consulting and publication and citation outcomes

The results from the endogenous switching models on research outcomes in the post-survey period are presented in Table 3.5 which shows the outcome equation for consulting-active and non-consulting-active academics separately. Exiting academics, i.e. those that do not publish in the post-survey period, are excluded from these models as we are only interested in the productivity effects for those who remained research active. When we run these models inclusive of those that exit, estimated coefficients are naturally larger because they capture the “exit” effects from consulting as reflected in more zeros in the outcome variable.

Our results suggest that consulting does not have a significant effect on publication numbers (model 1), which is contrary to prior studies by Mitchell and Rebne (1995) and Rentocchini et al. (2014) who found a positive or negative effect respectively. Consulting is however associated with fewer citations (model 3), an effect that stems from public sector consulting (model 4). The coefficient of public sector consulting is -0.014 and the squared term is positive, though very small, indicating a diminishing negative impact at higher consulting shares rather than a positive one. The AME of public consulting is still negative at -0.013 (s.e. = 0.004) which indicates that an increase in the public consulting share by 10 points leads to an average loss of 13% of field-weighted citations per publication. The coefficient of private sector consulting is also negative, but much smaller in magnitude and insignificant. The AME is 0.010, but only significant at the 10% level. These results suggest that public sector consulting could allow for fewer research spillovers as it primarily requires the preparation of reviews and commissioned reports that may result in publications of only little academic relevance (Salter 1988; Jasanoff 1990). While we find no negative effect for private sector consulting, we also do not confirm the positive effect for low levels reported in prior research (Mitchell and Rebne 1995; Rentocchini et al. 2014).

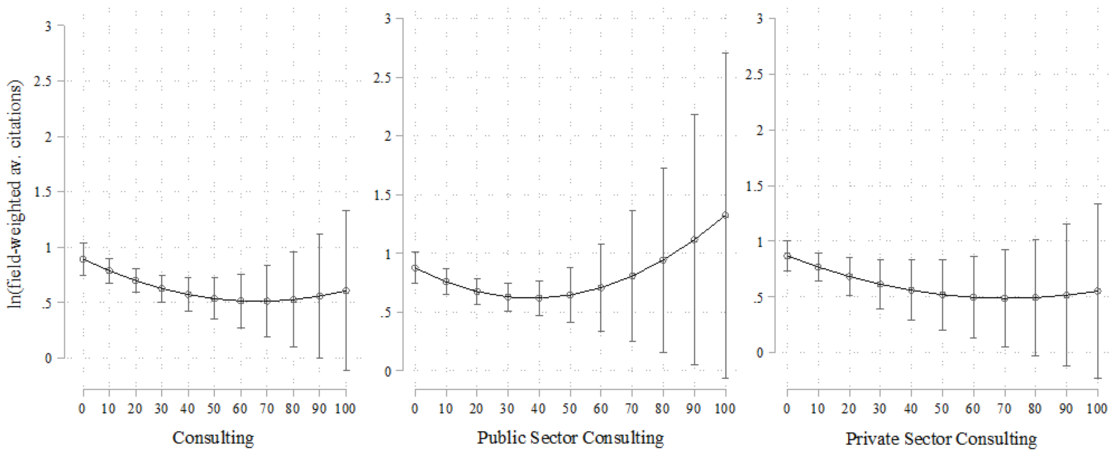
Again, the effect is likely not linear and the marginal effect may depend on the intensity of consulting. Figure 3.4 therefore depicts predicted values of field-weighted average citations as outcome variable over the consulting time-share range. The slope of the curve illustrates the marginal effect of consulting at different levels of consulting (MERs). Here we see that an increase in public consulting from zero to 5% implies a decline in the predicted logged number of field-weighted citations from 0.88 to 0.81. In non-weighted and non-logged terms, the same increase in consulting results in a decline from 11.5 to 9.9 average citations per paper, i.e. to the loss of 1.6 citations per paper which corresponds to about 19% of the sample median. At higher consulting shares the marginal effect of public consulting becomes smaller and eventually insignificant. For private sector consulting the effects are insignificant for the full range of consulting shares.

Table 3.5: Estimation results from endogenous switching models on research outcomes (without "exits")

	ln(field-weighted publications$_{2009-2013}$)		**ln(field-weighted publications$_{2009-2013}$)**		**ln(field-weighted av. citations$_{2009-2013}$)**		**ln(field-weighted av. citations$_{2009-2013}$)**	
	no consulting	consulting	no consulting	consulting	no consulting	consulting	no consulting	consulting
consulting		<-0.001				-0.011 ***		
		(0.004)				(0.004)		
consulting²		<-0.001				<0.001		
		(<0.001				(<0.001)		
public consulting				-0.006				-0.014 ***
				(0.005)				(0.005)
public consulting²				<0.001				<0.001 *
				(<0.00				(<0.001)
private consulting				-0.002				-0.011 *
				(0.005)				(0.006)
private consulting²				<-				<0.001
				(<0.00				(<0.001)
age	0.043 **	0.076 ***	0.043 **	0.077 ***	-0.006	0.030	-0.006	0.031
	(0.019)	(0.027)	(0.019)	(0.027)	(0.028)	(0.022)	(0.028)	(0.022)
age^2	-0.001 ***	-0.001 ***	-0.001 ***	-0.001 ***	<0.001	<-0.001	<0.001	<-0.001
	(<0.001)	(<0.001	(<0.000	(<0.00	(<0.001)	(<0.001)	(<0.001)	(<0.001)
ln(average number of co-authors)	0.055	0.029	0.055	0.026	0.078 **	-0.071	0.078 **	-0.071
	(0.047)	(0.050)	(0.047)	(0.050)	(0.031)	(0.057)	(0.031)	(0.057)
ln(field-weighted pubs)$_{2002-2008}$ / ln(field-weighted	0.568 ***	0.680 ***	0.568 ***	0.683 ***	0.506 ***	0.490 ***	0.506 ***	0.485 ***
	(0.042)	(0.043)	(0.042)	(0.043)	(0.071)	(0.062)	(0.071)	(0.062)
no_publication_d / no_avcit_d	0.190 **	0.216 **	0.190 **	0.219 **	0.277 ***	0.029	0.277 ***	0.029
	(0.093)	(0.095)	(0.093)	(0.095)	(0.083)	(0.124)	(0.083)	(0.124)
ln(industry funding)$_{2002-2006}$	-0.120	-0.018	-0.120	-0.002	-0.134	0.309 *	-0.134	0.311 **
	(0.166)	(0.111)	(0.166)	(0.120)	(0.175)	(0.159)	(0.176)	(0.156)
ln(public funding)$_{2002-2006}$	0.069	<0.001	0.069	0.001	0.092	-0.061	0.092	-0.060
	(0.057)	(0.074)	(0.057)	(0.076)	(0.067)	(0.077)	(0.067)	(0.075)
collaborative reach$_{2002-2008}$	0.028 **	0.019	0.028 **	0.022	0.016	0.003	0.016	0.004
	(0.012)	(0.016)	(0.012)	(0.016)	(0.016)	(0.017)	(0.016)	(0.017)
ln(patents$_{pre2009}$)	0.007	0.070 **	0.007	0.067 **	-0.027	-0.033	-0.027	-0.034
	(0.027)	(0.029)	(0.027)	(0.029)	(0.024)	(0.029)	(0.024)	(0.029)
Log pseudolikelihood	-643.51		-642.52		-776.96		-776.52	
Wald test of indep. equations	4.16		3.94		7.95**		7.74**	
α^{ρ} (consulting = 0)	-0.074 (0.175)		-0.071 (0.173)		0.111 (0.234)		0.110 (0.236)	
α^{ρ} (consulting = 1)	-0.398 (0.186)**		-0.387 (0.187)**		-0.358 (0.126)***		-0.356 (0.127)***	

Notes: N = 784. *** (**,*) indicate a significance level of 1% (5%, 10%). All models contain a gender dummy, a dummy variable for co-authorship with industry, ln(peer group size), international visibility and rank dummies as well as a variable indicating university affiliation and a constant. Coefficients presented; robust standard errors in parenthesis below. For unweighted publication and citation variables the signs and significance levels are similar. Outcome variables and logged controls are calculated as the natural logarithm of the variable plus one. First stage estimation results and results for the full sample available upon request.

Figure 3.3: Predictive margins with 95% confidence intervals for field-weighted av. citations per publication (without "exits", 784 obs.)



Notes: Predicted outcome variable depicted in the y-axis and consulting share on the x-axis. Predictive margins are calculated at the deciles of the consulting time-share distribution. The slope of the predictive margins curve represents the marginal change in the predicted value of the outcome variable for a change in the consulting time-share, i.e. the marginal effect.

Looking at the MERs of moderators in Figure 3.5, we see that fewer field-weighted average citations are observed for most of the range of consulting time-shares for all academic ranks (detailed regression results available upon request). For assistant and full professors the effect reverses (i.e. turns positive) in the top percentiles of the consulting time-share distribution. Instead, the AME is -0.028 (s.e. = 0.017) for junior researchers and -0.038 (s.e. = 0.011) for assistant professors and thus substantially more negative than the overall effect and particularly compared to the -0.016 (s.e. = 0.008) for full professors. This confirms our expectation that those with fewer consulting opportunities will be less able to generate positive spillovers for their research.

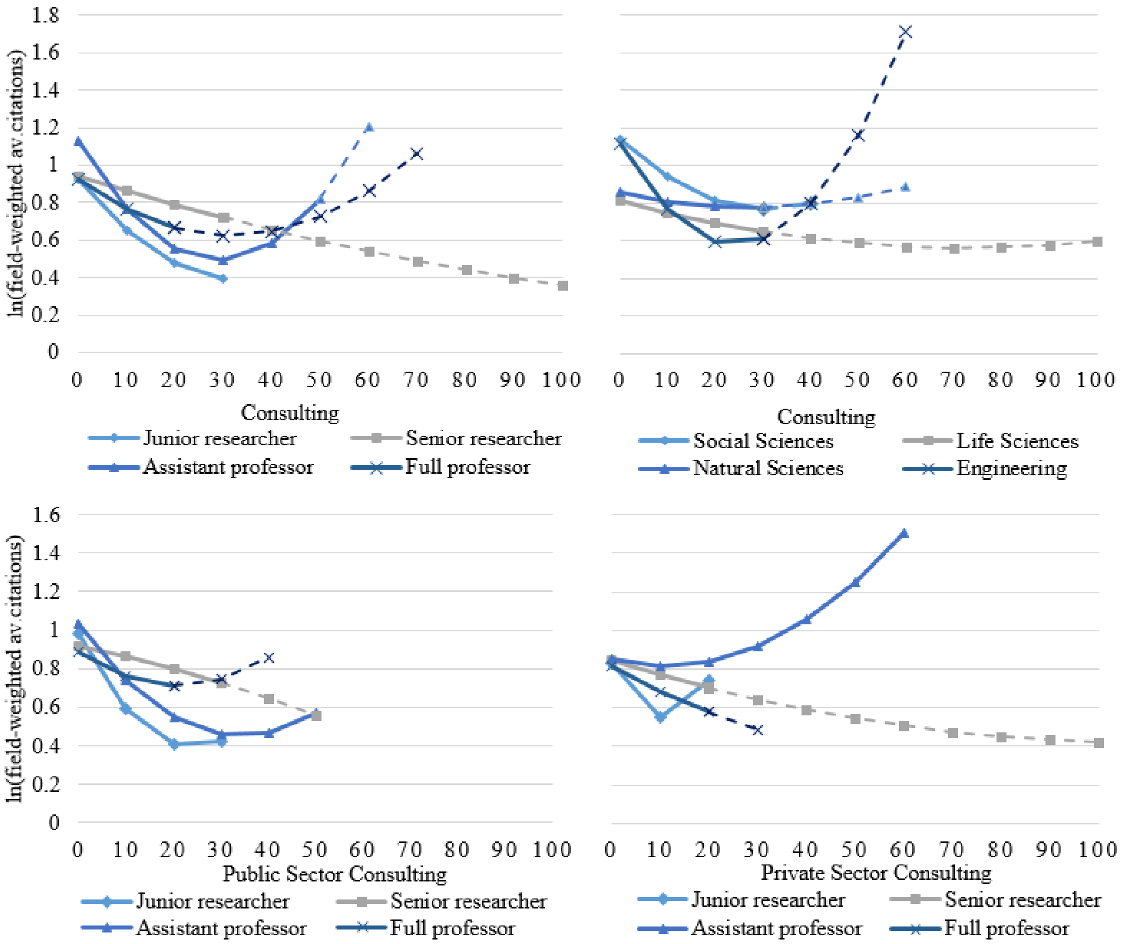
The lower part of Figure 3.5 differentiates the effects by type of consulting. Again we see that for all academic ranks the impact of public consulting is negative for most of the observed consulting time-share distribution. The impact is strongest for junior researchers [AME = -0.044 (s.e. = 0.014)] and assistant professors [AME = -0.032 (s.e. = 0.011)] and more modest for full professors [AME -0.015 (s.e. = 0.006)] and senior researchers [-0.005 (s.e. = 0.010)]. In terms of non-field weighted citations this implies that for junior researchers an increase from zero to 10% consulting time-share leaves them with about 2.3 fewer citations per paper which corresponds to 31% of the median value for junior researchers.

Some assistant professors who engage heavily in consulting efforts with the public sector, however, benefit and receive relatively more citations than those at lower consulting time-shares, but still not more than those not active in consulting at all. Private sector consulting has no negative and even a positive effect on outcomes of assistant professors, but a negative effect for full professors [AME = -0.018 (s.e. = 0.010)] and, at lower consulting shares, also for junior researchers [AME = -0.049 (s.e. = 0.023)]. These mixed effects may explain the overall only very weakly significant effect from private sector consulting.

Differentiating by disciplinary field (top right of Figure 3.5), we see a continuous negative effect only for the life sciences, which is significantly negative up to a time-share of 50%. The marginal effects for natural sciences are negative up to 20% consulting time-share. In engineering, on the other hand, the marginal effects are initially negative and the curve has the steepest slope of all subject areas, but marginal effects become positive and significant for values above 30%, i.e. above the 90^{th} percentile in this field. The AME [-0.031, (s.e. = 0.008)] is still negative and sizable for engineering. Our results only partially confirm Rentocchini et al. (2014) who find a negative effect of paid consulting in science and engineering but not in medical sciences and social sciences. They also find the strongest effect at high engagement levels, which is contrary to our findings that show the steepest slopes in the middle-ranges.

We further see from the models presented in Table 3.5 that publication and citation performance is highly path-dependent. The pre-sample mean is positive, highly significant and the coefficients are similar in size for both consulting-active and non-consulting-active academics. We also find that publication output is larger for older academics and for professors. We do not observe differences between men and women regarding their publishing when we use field-weighted publication counts. Scientific attributes such as collaborative reach and international visibility are also positively associated with publication output. We also find that publication numbers are lower for university academics, who have teaching obligations unlike most academics at PROs, whereas average citations do not differ. Patents are positively associated with field-weighted publication numbers for consulting-active academics only. Note that the correlation coefficient α^{ρ} is negative and significant only for the correlation between the consulting equation and the outcome equation for consulting-active individuals. This suggests that individuals who engage in consulting publish fewer articles and receive fewer citations than a random individual from the sample would have published. Instead, those not engaged in consulting do no better or worse than the sample average. The likelihood-ratio test for joint independence of the three equations, however, is not significant in the publication count models where we exclude "exited" individuals suggesting that consulting and publication equations are not jointly determined. It should be noted, however, that the test is significant in models that include those that "exit". In other words, much of the endogeneity in terms of two-way causation is taken out of the model by considering only those who remain research active.

Figure 3.4: Predictive margins for field weighted av. citations per publication by rank and field (without "exits", 784 obs.)



Notes: Predictive margins are only shown for the range of consulting values where the margins are significant at least at 10% confidence level and only within the relevant observed consulting intensity range for the respective group. Predicted outcome variable depicted in the y-axis and consulting share on the x-axis. Predictive margins are calculated at the deciles of the consulting time-share distribution. The slope of the predictive margins curve represents the marginal change in the predicted value of the outcome variable for a change in the consulting time-share, i.e. the marginal effect. Dashed lines indicate values beyond the 95th percentile of the consulting share distribution.

3.5 Conclusions and implications

Our study contributes to the literature on academic consulting and its impact on research and research dissemination. Investigating the effect of public and private sector consulting activities on exit from publishing and publication performance in a sample of academics at universities and public research organizations in Germany, we find that, especially in the case of private sector consulting, a higher share of time devoted to consulting increases the probability of exit from academic work. At higher consulting time-shares this effect is strong for lower rank (pre-PhD) researchers, but also for faculty in permanent positions (full professors). Public sector consulting also affects exit probability, but only at relatively high consulting shares. The positive relationship between consulting and exit from publishing is more pronounced in the social sciences and engineering than in the natural sciences and life sciences. This is consistent with the observations that the public-private wage gap is particularly high for engineers while opportunities for taking up new responsibilities outside academia, or external demand, may be particularly high for those from the social sciences and engineering.

Results for academics who remain research-active show that consulting does not further reduce their ex-post research performance in terms of publication numbers. This result thus does not confirm concerns related to a potential detrimental effect of consulting on research disclosure as we do not find a decline in overall publication numbers. However, in the case of public consulting, we see lower average citations per paper in the ex-post period. Public sector consulting, likely requires participation in expert committees and boards of advisors which comes with the preparation of reviews and commissioned reports and thus entails work aspects that may create few financial and other positive spillovers for academic research. Quantitatively we show that an increase in public consulting by ten percentage points implies a loss of up to 31% of citations per paper. While this price of consulting is not paid by researchers from all ranks, it should be noted that the decline is most pronounced for junior researchers. Considering, that typically junior researchers are still seeking to obtain permanent positions this puts them at a potential disadvantage on the academic labor market compared to their peers. It may also have a longer term impact on their research paths.

The finding that private sector consulting, instead, does not impact research output, once we exclude non-publishing academics, suggests that it may be closer to the knowledge frontier and may therefore create more research spillovers which offset some of the negative trade-offs. Still, negative effects are observed for junior researchers, suggesting that they may lack the experience to realize such spillovers.

Disciplinary field differences exist in the prevalence of consulting to different sectors, but less so in terms of the impact of consulting. While in the natural sciences consulting has little impact on citations, in the social and life sciences and in engineering higher engagement in consulting is associated with fewer citations per publication. In engineering the marginal effect turns, however, positive for consulting time-shares above the 90^{th} percentile. This indicates that at the higher end consulting can create positive spillovers in more applied fields of science that apply academic knowledge to real-world problems. Thus, for highly engaged academics in engineering there seems to be a prize for consulting.

Our findings have important implications for research institutions and policy. First, for academics in earlier stages of their academic career and also for senior academic staff, consulting activities may pave the way for alternative career paths or activities outside academic research, as indicated by an exit from academic publishing. Training and institutional consulting support for junior academic staff could thus have the potential to open up career options outside academic research. Professors and research group leaders may engage junior researchers in consulting work to broaden their profile and to point to career opportunities outside academe. The provision of alternative options is important as not all those trained in academia are able to remain there (e.g. Stephan 2012; Hottenrott and Lawson 2017). However, encouraging external consulting could also lead to a brain drain at both junior and senior levels if academics cease to focus on scientific research relevant for the scientific community. This may also have detrimental career effects for those young researchers seeking an academic career path.

Second, our selection equations show that academics that engage in consulting are on average involved in more grant acquisition and are highly connected. They may therefore serve as important knowledge brokers with external organizations, leveraging additional income for their institution while providing advice. While this may come at the price of lower quality research output or the exit from academic publishing, it may contribute to a division of labor within the academic institution that allows for different work patterns amongst academics. Universities may therefore selectively encourage specific academics to act as such knowledge brokers.

Third, policies (e.g. promotion requirements) to engage *all* academics to interact with external organizations may have negative consequences for academic research. In particular, explicit or implicit obligations to take on consulting roles should not exist. We find that academics that do not engage in consulting are often less focused on external interactions in general and pursue research that attracts more citations. Such individuals may, as a result of engagement policies, have their time diverted from their research efforts to the detriment of their research. Eventually, the results suggest that a one-size-fits-all rule for managing consulting activities of researchers at universities or PROs will not work best, but that universities may be advised to arrange disciplinary and rank specific rules. Specifically, the consulting activities of junior researchers need to be carefully managed.

Overall, the benefits from academic consulting likely outweigh the costs in terms of research output. For example, Cohen et al. (2002) report that 32% of surveyed US firms consider consulting an important mechanism to gain insights into academic research. This figure is higher than for other forms of knowledge transfer such as contract research, patents or personnel exchanges. In the case of public consulting, Haucap and Thomas (2014) find in a survey of more than 300 civil servants and politicians in Germany that more than 70% of users of academic knowledge consider expert reports and personal communication with academics as helpful or very helpful for their work, making consulting more important than academic publications. Thus, while we do find some negative effects on research quality as measured through citations, we cannot conclude that the price of consulting is high compared to the likely benefits for private sector firms and public sector agents.

Despite all efforts, the study is not without some limitations. First, we do not observe consulting activity over time. Individuals may undergo different phases in their career in which they are more or less consulting active. The balance between these periods could be pivotal to understand the full effects of consulting engagement. Second, some individual level unobserved heterogeneity might remain despite our attempts to capture these econometrically. Thus, longitudinal treatment effects analysis might be used in future research to test for the observed cross-sectional patterns. Finally, some limitations arise in terms of generalizability of our results to the overall population of researchers in Germany and to the population of academics in general. Individual wage levels, specific salary schemes or contracts may determine the attractiveness of consulting. Further the division of public research into universities, universities of applied sciences and public research organization without teaching mission in Germany and the mobility of researchers between these institutions may have implications for the results. We therefore encourage further research on academic consulting especially regarding its role for inter-sector mobility of academics and for the evolution of career paths. Moreover, while we considered time shares rather than monetary rewards for consulting, it would be desirable to better understand the link between remuneration and the effects of consulting on other academic activities. While well paid consulting that is informed by research may increase the academics' institutional research budget through follow-up research contracts and therefore facilitate growth and productivity of the research group, consulting activities that result in private income may be more prone to lead to a brain drain from academic work. It seems therefore crucial to further study the contractual mechanisms in future work.

References

Abreu, M. and Grinevich, V. (2013), Academic Interactions with Private, Public and Not-for-Profit Organisations: The Known Unknowns. In: Ferreira, J., Raposo, M., Rutten, R. and Varga, A. (eds) *Cooperation, Clusters, and Knowledge Transfer. Advances in Spatial Science*. Berlin Heidelberg: Springer, 181-206.

Agarwal, R and Ohyama, A. (2013), Industry or Academia, Basic or Applied? Career Choices and Earnings Trajectories of Scientists, *Management Science 59*, 950–970.

Amara, N., Landry, R. and Halilem, N. (2013), Faculty Consulting in Natural Sciences and -Engineering: Between Formal and Informal Knowledge Transfer, *Higher Education* 65(3), 359–384.

Azoulay, P., W. Ding and Stuart, T. (2009), The Impact of Academic Patenting on the Rate, Quality, and Direction of (Public) Research, *The Journal of Industrial Economics*, 57(4), 637–676.

Balsmeier, B. and Pellens, M. (2014), Who Makes, who Breaks: Which Scientists Stay in Academe? *Economics Letters* 122(2), 229–232.

Balsmeier, B. and Pellens, M. (2016), How much does it cost to be a scientist? *The Journal of Technology Transfer* 41(3), 469–505.

Banal-Estañol, A., Jofre-Bonet, M. and Lawson, C. (2015), The Double-Edged Sword of Industry Collaboration: Evidence From Engineering Academics in the UK, *Research Policy* 44, 1160-1175.

Bekkers, R., Bodas Freitas, I.M. (2008), Analysing knowledge transfer channels between universities and industry: to what degree do sectors also matter? *Research Policy* 37 (10), 1837–1853.

Bianchini, S., Lissoni, F., Pezzoni, M. and Zirulia, L. (2016), The economics of research, consulting, and teaching quality: theory and evidence from a technical university, *Economics of Innovation and New Technology* 25, 668-691.

Bischof, D. (2016), BLINDSCHEMES: Stata module to provide graph schemes sensitive to color vision deficiency. Statistical Software Components S458251. Chestnut Hill MA: Boston College.

Blumenthal, D., Campbell, E.G., and Causino, N. (1996), Participation of Life-Science Faculty in Research Relationships with Industry, *New England Journal of Medicine* 335, 1734-1739.

Boardman, C.P. and Ponomariov, B.L. (2009), University Researchers Working with Private Companies, *Technovation* 29(2), 142–153.

Bodas Freitas, I.M., Geuna, A. and Rossi, F. (2013), Finding the right partners: Institutional and personal modes of governance of university–industry interactions, *Research Policy* 42 (1), 50-62

Boyer, C.M. and Lewis, D.R. (1984), Faculty Consulting : Responsibility or Promiscuity?, *The Journal of Higher Education* 55(5), 637–659.

Boyne, G.A. (2002), Public and Private Management: What's the Difference?, *Journal of Management Studies* 39, 97-122.

Breschi, S., Lissoni, S., Montobbio, F. (2007). The scientific productivity of academic inventors: new evidence from Italian data, *Economics of Innovation and New Technol*ogy 16 (2), 101–118.

Buenstorf, G. (2009), Is Commercialization Good or Bad for Science? Individual-Level Evidence From the Max Planck Society, *Research Policy* 38, 281–292.

BUWIN (2017), *Bundesbericht Wissenschaftlicher Nachwuchs 2017*, Institut für Innovation und Technik (iit), Berlin.

Cohen, W.M., Nelson, R.R. and Walsh, J.P. (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science* 48, 1–23.

Czarnitzki, D. and Toole, A. (2010), Is There a Trade-off between Academic Research and Faculty Entrepreneurship? Evidence from US NIH Supported Biomedical Researchers, *Economics of Innovation and New Technology* 19(5), 505-520.

Czarnitzki, D., Grimpe, C. and Toole, A. (2015), Delay and Secrecy: Does Industry Sponsorship Jeopardize Disclosure of Academic Research?, *Industrial and Corporate Change* 24(1), 251-279.

D'Este, P. and Patel, P. (2007), University-industry linkages in the UK: What are the factors underlying the variety of interactions with industry? *Research Policy* 36(9), 1295-1313.

D'Este, P., Rentocchini, F., Grimaldi, R. and Manjarrés-Henríquez, L. (2013), The Relationship Between Research Funding and Academic Consulting: An Empirical Investigation in the Spanish Context, *Technological Forecasting and Social Change* 80(8), 1535–1545.

De Luca, G. (2008), SNP and SML Estimation of Univariate and Bivariate Binary-Choice Models, *Stata Journal* 8, 190–220.

DFG (2015), *Förderatlas 2015 - Kennzahlen zur öffentlich finanzierten Forschung in Deutschland*, 1. Auflage, Weinheim: Wiley-VCH.

Emran, M.S. and Hou, Z. (2013), Access to Markets and Rural Poverty: evidence from Household Consumption in China, *The Review of Economics and Statistics* 95(2), 682-697.

Erk, D. and Schmidt, M. (2014), Nebenverdienst von Professoren: Gelehrte Geschäftsleute, http://www.zeit.de/2014/29/nebenverdienst-professoren-hochschule (retrieved on 26.04.2010).

Etzkowitz, H. and Webster, A. (1998), Entrepreneurial Science: The Second Academic Revolution, In H. Etzkowitz, A. Webster, and P. Healey (eds.), Capitalizing Knowledge: New Intersections of Industry and Academia (pp. 21-46), Albany: *State University of New York Press*.

Fabrizio, K. and DiMinin, A. (2008), Commercializing the Laboratory: Faculty Patenting and the Open Science Environment, *Research Policy* 37, 914–931.

Florida, R. and W.M. Cohen (1999), Engine or Infrastructure? The University's Role in Economic Development, In: L.M. Branscomb, F. Kodama and R. Florida (Eds.), Industrializing Knowledge: University-Industry Linkages in Japan and the United States (pp. 589–610), Cambridge: *MIT Press*.

Ginther, D. K. and Kahn, S. (2004), Women in Economics: Moving up or Falling off the Academic Career Ladder?, *Journal of Economic Perspectives* 18(3), 193–214.

Grimpe, C. and Fier, H. (2010), Informal University Technology Transfer: a Comparison Between the United States and Germany, *The Journal of Technology Transfer* 35, 637–650.

Grimpe, C. and Hussinger, K. (2013), Formal and Informal Knowledge and Technology Transfer from Academia to Industry: Complementarity Effects and Innovation Performance, *Industry and Innovation* 20, 683–700.

Gulbrandsen, M. and Smeby, J.C. (2005), Industry Funding and University Professors' Research Performance, *Research Policy* 34, 932–950.

Haucap, J. and Thomas, T. (2014), Wissenschaftliche Politikberatung: Erreicht der Rat von Ökonomen Politik und Öffentlichkeit?, *Wirtschaftsdienst* 94(3), 180–186.

Haucap, J. and Mödl, M. (2013), Zum Verhältnis von Spitzenforschung und Politikberatung. Eine empirische Analyse vor dem Hintergrund des Ökonomenstreits, *Perspektiven der Wirtschaftspolitik* 14(3-4), 346–378.

HEFCE (2017) *Higher Education-Business and Community Interaction survey 2015-16*. Higher Education Funding Council for England (HEFCE) No. 2017/23.

Hottenrott, H. (2012), *The Role of Research Orientation for Attracting Competitive Research Funding*. In: Audretsch/Lehmann/Link/Starnecker (Hrsg.), International Handbook Series on Entrepreneurship, 35-49: Springer.

Hottenrott, H. and Thorwarth, S. (2011), Industry Funding of University Research and Scientific Productivity, *Kyklos* 64, 534–555.

Hottenrott, H. and Lawson, C. (2014), Research Grants, Sources of Ideas and the Effects on Academic Research, *Economics of Innovation and New Technology* 23, 109–133.

Hottenrott, H. and Lawson, C. (2017), Flying the Nest: How the Home Department Shapes Researchers' Career Paths, *Studies in Higher Education*, 42, 1091-1109.

Hughes, A., Lawson, C., Salter, A., Kitson, M. with Bullock, A. and Hughes, R.B. (2016), The Changing State of Knowledge Exchange: UK Academic Interactions with External Organisations 2005 -2015, NCUB, London.

Jacobson, N., Butterill, D. and Goering, P. (2005), Consulting as a strategy for knowledge transfer, *Milbank Quarterly*, 83(2), 299-321.

Jasanoff, S. (1990), The Fifth Branch. Science Advisers as Policymakers. Cambridge, MA: *Harvard University Press*.

Jensen, R., Thursby, J. and Thursby, M.C. (2010), University-Industry Spillovers, Government Funding, and Industrial Consulting, NBER Working Paper No. 15732, Cambridge, M.A.

Klofsten, M. and Jones-Evan,s D. (2000), Comparing Academic Entrepreneurship in Europe– the Case of Sweden and Ireland, *Small Business Economics* 14(4), 299-309.

Krimsky, K. (2003), Science in the Private Interest: Has the Lure of Profits Corrupted the Virtue of Biomedical Research?, Lanham: *Rowman & Littlefield*.

Landry, R., Saïhi, M., Amara, N. and Ouimet, M. (2010), Evidence on how Academics Manage Their Portfolio of Knowledge Transfer Activities, *Research Policy* 39, 1387–1403.

Lee, Y.S. (2000), The Sustainability of University-Industry Research Collaboration: An Empirical Assessment. *The Journal of Technology Transfer,* 25, 111-133.

Lewbel, A. (2012), Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models, *Journal of Business & Economic Statistics* 30(1), 67-80.

Link, A.N., Siegel, D.S. and Bozeman, B. (2007), An Empirical Analysis of the Propensity of Academics to Engage in Informal University Technology Transfer, *Industrial and Corporate Change* 16(4), 641–655.

Lokshin, M., and Z. Sajaia (2004), Maximum Likelihood Estimation of Endogenous Switching Regression Models, *Stata Journal* 4, 282-289.

Louis, K.S., Blumenthal, D., Gluck, M.E. and Stoto, M.A. (1989), Entrepreneurs in Academe: An Exploration of Behaviours Among Life Scientists, *Administrative Science Quarterly* 34(1), 110–131.

Lowen, R. (1996). *Creating the Cold War University: The Transformation of Stanford*, University of California Press, Berkeley, CA.

Mairesse, J. and Pezzoni, M. (2015), Does Gender Affect Scientific Productivity? A Critical Review of the Empirical Evidence and a Panel Data Econometric Analysis for French Physicists. *Revue économique* 66(1), 65-113.

Manjarrés-Henríquez, L., Gutiérrez-Gracia, A., Carrión-García, A., and Vega-Jurado, J. (2009), The effects of university–industry relationships and academic research on scientific performance: Synergy or substitution?, *Research in Higher Education* 50(8), 795-811.

Mitchell, J.E. and Rebne, D.S. (1995), Nonlinear Effects of Teaching and Consulting on Academic Research Productivity, *Socio-Economic Planning Sciences* 29(1), 47–57.

Muscio, A., D. Quaglione and Vallanti, G. (2013), Does government funding complement or substitute private research funding to universities?, *Research Policy* 42, 63– 75.

OECD (2015), Scientific Advice for Policy Making: The Role and Responsibility of Expert Bodies and Individual Scientists, OECD Science, Technology and Industry Policy Papers, No. 21, Paris: *OECD Publishing*.

Perkmann, M. and Walsh, K. (2008), Engaging the Scholar: Three Types of Academic Consulting and their Impact on Universities and Industry, *Research Policy* 37(10), 1884–1891.

Perkmann, M. and Walsh, K. (2009), The two Faces of Collaboration: Impacts of University-Industry Relations on Public Research, *Industrial and Corporate Change* 18(6), 1033–1065.

Perkmann, M., Fini, R., Ross, J.M, Salter, A. Silvestri, C. and Tartari, V. (2015), Accounting for universities’ impact: using augmented data to measure academic engagement and commercialization by academic scientists, *Research Evaluation* 24, 380–391.

Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., Fini, R., Geuna, A., Grimaldi, R., Hughes, A., Kitson, M., Krabel, S., Llerena, P., Lissoni, F., Salter, A. and Sobrero, M. (2013), Academic Engagement and Commercialisation: A Review of the Literature on University-Industry Relations, *Research Policy* 42(2), 423-442.

Rebne, D.S. (1989), Faculty Consulting and Scientific Knowledge: A Traditional University-Industry Linkage, *Educational Administration Quarterly* 25, 338–357.

Rentocchini, F., Manjarrés-Henrìquez, L., D'Este, P. and Grimaldi, R. (2014), The Relationship Between Academic Consulting and Research Performance: Evidence From Five Spanish Universities, *International Journal of Industrial Organization* 32(1), 70–83.

Roach, M. and H. Sauermann (2010), A Taste for Science? PhD Scientists' Academic Orientation and Self-Selection into Research Careers in Industry, *Research Policy* 39(3), 422–434.

Salter, L. (1988), *Mandated Science: Science and Scientists in the Making of Standards*. Kluwer Academic Publishers, Dordrecht, Holland.

Shimshoni, D. (1970), The Mobile Scientist in the American Instrument Industry, *Minerva* 8, 59–89.

Stephan, P.E. (2012), *How Economics Shapes Science*, Cambridge, MA: Harvard University Press.

Thursby, M., J. Thursby and Gupta-Mukherjee, S. (2007), Are There Real Effects of Licensing on Academic Research? A Life Cycle View, *Journal of Economic Behavior & Organization* 63(4), 577-598.

Toole, A.A. and Czarnitzki, D. (2010), Commercializing Science: Is There a University "Brain Drain" From Academic Entrepreneurship?, *Management Science* 56(9), 1599-1614.

van de Ven, W. and van Pragg, B. (1981), The demand for deductibles in private health insurance: A probit model with sample selection, *Journal of Econometrics* 17, 229–252.

van Looy, B., Callaert, J. and Debackere, K. (2006), Publication and Patent Behavior of Academic Researchers: Conflicting, Reinforcing or Merely Co-Existing? *Research Policy* 35(4), 596–608.

Vavakova, B. (1998), The new Social Contract Between Governments, Universities and Society: has the old one Failed?, *Minerva* 36(3), 209-228.

Wooldridge, J. (2002), *Econometric Analysis of Cross Section And Panel Data*, The MIT Press, Boston: MA.

Zucker, L. G., Darby, M. R. and Torero, M. (2002), Labor Mobility from Academe to Commerce, *Journal of Labor Economics* 20(3), 629–660.

Appendix

Supplementary descriptive statistics

Table 3.6: Correlation matrix of covariates (n = 951)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	age	1															
2	female	-0.123*	1														
3	rank	0.212*	-0.069	1													
4	publications$_{2002-2008}$	0.085*	-0.085*	0.168*	1												
5	average citations$_{2002-2008}$	-0.105*	-0.060	0.029	0.266*	1											
6	av. number of coauthors	0.036	-0.035	0.003	0.276*	0.084	1										
7	collaborative reach	-0.127*	-0.087*	0.048	0.243*	0.257*	0.067	1									
8	international visibility	-0.044	-0.028	0.013	0.091*	0.122*	-0.018	0.136*	1								
9	industry funding	0.078	-0.103*	0.049	0.052	-0.059	-0.021	0.136*	-0.037	1							
10	public funding	0.032	-0.049	0.132*	0.237*	0.122*	0.202*	0.270*	0.071	0.374*	1						
11	peer group size	0.010	-0.056	-0.035	0.043	0.132*	0.092	0.069	-0.030	0.048	0.154*	1					
12	university	0.098*	0.013	0.623*	0.045	-0.069	-0.048	-0.062	-0.015	-0.006	0.029	-0.080	1				
13	patents$_{pre2009}$	0.062	-0.059	-0.020	0.078	-0.008	-0.014	0.056	-0.009	0.337*	0.160*	0.014	-0.079	1			
14	firm	0.118*	-0.080	0.128*	0.087*	-0.034	-0.018	0.052	-0.004	0.230*	0.244*	0.048	0.084*	0.151*	1		
15	tech transfer industry	0.029	-0.114*	0.036	0.072	-0.040	-0.053	0.117*	-0.006	0.328*	0.236*	0.089*	-0.041	0.212*	0.312*	1	
16	co-authorship industry	0.027	-0.101*	0.049	0.093*	-0.053	-0.020	0.083	0.061	0.330*	0.183*	0.012	0.040	0.190*	0.224*	0.368*	1
17	regio skills	0.042	-0.013	0.072	-0.007	-0.014	0.080	-0.013	0.010	0.042	0.035	0.038	0.087	0.072	0.031	0.065	0.041

Notes: Rank is the ordinal version of the rank dummies (rank=1: Junior Researcher, rank=2: Senior Researcher, rank=3: Assistant Professor, rank=4: Full Professor). * indicates a significance level of at least 1%.

Table 3.7: Sample versus population distribution

	Sample share academic scientists (%)	Population share academic scientists (%)
Institution type (university vs. PRO)	58.57	67.6
Gender (female vs. male)	15.04	31.5
Field:		
natural sciences	30.7	31.1
engineering	18.51	19.3
life sciences	29.86	25.2
social sciences and humanities	20.93	24.4
Age cohort:		
less than 35	2.1	34.2
35 and less than 45	28.6	26.6
45 and less than 55	43.64	18.2
55 and less than 65	21.24	15.1
greater than 65	4.42	5.8

Note: Population data obtained from the data base of the German Federal Statistical Office (DESTATIS).

Table 3.8: Academics' division of time (in % of total time at work)

		Grant-funded research	Research (block funded)	Teaching	Admin.	Public sector consulting	Private sector
	obs.	mean (s.d.)	mean (s.d.)	mean (s.d.)	mean (s.d.)	mean (s.d.)	mean (s.d.)
Full sample	**951**	**32.1 (22.7)**	**20.0 (20.9)**	**21.5 (16.6)**	**21.1 (16.4)**	**3.1 (8.0)**	**2.2 (6.2)**
By rank							
junior Researcher	90	31.6 (31.1)	35.5 (32.7)	8.9 (15.6)	16.9 (21.4)	5.6 (16.2)	1.4 (3.5)
senior researcher	243	40.7 (24.6)	22.5 (22.3)	9.5 (10.5)	21.1 (19.2)	3.0 (7.4)	3.2 (8.7)
assistant	107	37.8 (26.3)	22.1 (22.5)	21.5 (14.7)	15.1 (14.1)	2.1 (6.6)	1.5 (7.0)
full professor	511	27.0 (17.1)	15.6 (14.7)	29.4 (14.7)	23.2 (13.7)	2.8 (6.0)	2.1 (4.7)
By discipline							
social Sciences	199	24.8 (23.8)	22.3 (20.1)	28.1 (18.6)	19.2 (15.1)	3.9 (7.7)	1.6 (5.1)
life Sciences	284	34.7 (22.2)	19.3 (21.9)	18.5 (14.2)	22.0 (16.3)	3.5 (7.9)	2.0 (7.0)
natural Sciences	292	32.0 (22.1)	22.4 (23.4)	21.0 (16.7)	21.5 (17.9)	1.9 (5.8)	1.2 (3.3)
engineering	176	36.5 (21.3)	14.4 (13.2)	19.5 (16.0)	21.3 (15.2)	3.3 (10.8)	5.0 (8.5)
By consulting							
consulting inactive	537	33.8 (24.9)	22.5 (23.2)	22.7 (17.7)	21.0 (17.7)	0	0
consulting active	414	30.0 (19.2)	16.7 (16.7)	19.9 (15.0)	21.3 (14.5)	7.0 (10.9)	5.2 (8.6)
Pr(|T| > |t|)		***	***	**		***	***

Notes: *** (**,*) indicate a significance level of 1% (5%, 10%). Variable means presented. Standard deviations in parentheses.

Supplementary regression tables

Table 3.9: Results of probit and simultaneous probit models on private and public sector consulting (with inverse probability weights)

Model	1			2					
Dependent variable	**consulting** [yes / no]			**public consulting** [yes / no]			**private consulting** [yes / no]		
	df/dx	s.e.		df/dx	s.e.		df/dx	s.e.	
Moderators									
junior researcher				*Reference Category*					
senior researcher	-0.002	0.005		-0.052	0.006	***	0.011	0.005	**
assistant professor	-0.092	0.023	***	-0.117	0.022	***	-0.061	0.019	***
full professor	0.094	0.035	***	0.044	0.037		0.075	0.020	***
social sciences				*Reference Category*					
life sciences	-0.102	0.068		-0.109	0.071		0.040	0.013	***
natural sciences	-0.241	0.027	***	-0.212	0.051	***	-0.056	0.036	
engineering	-0.112	0.036	***	-0.134	0.030	***	0.045	0.045	
Controls									
age	0.002	0.001	**	0.004	0.001	***	0.001	0.001	*
female	-0.005	0.019		0.033	0.033		-0.053	0.026	**
field-weighted publications$_{2002\text{-}2008}$	0.006	0.003	**	0.006	0.004		0.008	0.003	***
field-weighted av. citations$_{2002\text{-}2008}$	-0.024	0.005	***	-0.027	0.006	***	-0.007	0.008	
collaborative reach$_{2002\text{-}2008}$	0.017	0.004	***	0.036	0.006	***	-0.005	0.010	
international visibility	0.032	0.045		0.004	0.031		0.095	0.061	
ln(industry funding)$_{2002\text{-}2006}$	0.265	0.114	**	-0.081	0.066		0.296	0.086	***
ln(public funding)$_{2002\text{-}2006}$	0.061	0.040		0.183	0.027	***	-0.084	0.026	***
ln(peer group size)	-0.014	0.009	*	-0.020	0.009	**	-0.005	0.012	
university	-0.110	0.030	***	-0.080	0.040	**	-0.058	0.013	***
ln(patents$_{pre2009}$)	-0.007	0.025		-0.050	0.014	***	0.012	0.011	
co-authorship industry	0.091	0.044	**	0.051	0.046		0.065	0.037	*
Exclusion restrictions									
regio skills	-0.006	0.002	**	-0.007	0.003	**	0.025	0.032	
firm	0.046	0.015	***	0.064	0.016	***	0.056	0.028	**
techtransfer industry	0.194	0.021	***	0.068	0.012	***	0.245	0.015	***
Log pseudolikelihood	-139,491.32			-226,887.20					
α^{ρ}	-			0.549 (0.052)***					

Notes: Number of observations = 951. Marginal effects at means. *** (**,*) indicate a significance level of 1% (5%, 10%). All models contain a constant. If we include unweighted publication and citation variables the signs and significance levels are similar.

Table 3.10: Results of Tobit and simultaneous Tobit models on private and public sector consulting

Model	**1**		**2**			
Dependent variable	**consulting** (time share)		**public consulting** (time share)		**private consulting** (time share)	
	coef.	s.e.	coef.	s.e.	coef.	s.e.
Moderators						
junior researcher			*Reference Category*			
senior researcher	-0.922 *	0.556	-2.245 ***	0.312	1.543 ***	0.222
assistant professor	-4.672 ***	0.680	-2.211 ***	0.514	1.136 ***	0.198
full professor	1.599	2.077	-1.273 **	0.580	1.316 ***	0.328
social sciences			*Reference Category*			
life sciences	-3.210	0.228	-0.969	0.590	0.244	0.580
natural sciences	-10.209 ***	2.371	-2.658 ***	0.413	-0.498	0.461
engineering	-2.467	2.950	-0.777	1.882	1.162	0.736
Controls						
age	0.073	0.061	0.015	0.030	-0.015	0.015
female	2.164 **	1.095	1.341 ***	0.514	0.266	0.246
field-weighted publications$_{2002\text{-}2008}$	0.206	0.137	0.046	0.060	-0.013	0.056
field-weighted av. citations$_{2002\text{-}2008}$	-1.190 **	0.504	-0.186 ***	0.061	-0.120	0.076
collaborative reach$_{2002\text{-}2008}$	-0.006	0.452	0.155	0.170	-0.093	0.092
international visibility	-7.937	5.383	-3.941	3.867	-2.203 ***	0.831
ln(industry funding)$_{2002\text{-}2006}$	9.752 *	5.142	-4.082	2.856	0.225 ***	0.088
ln(public funding)$_{2002\text{-}2006}$	1.786	2.774	1.989 *	1.169	-0.059 ***	0.019
ln(peer group size)	-0.607	0.502	0.007	0.141	-0.140	0.333
university	-7.786 ***	2.005	-2.685 ***	0.858	-1.478 ***	0.405
ln(patents$_{pre2009}$)	-1.853 *	1.024	-0.843 **	0.416	-0.477 ***	0.158
co-authorship industry	3.255	2.270	1.040	1.005	0.060	0.397
Exclusion restrictions						
regio skills	-0.325 ***	0.110	-0.089 *	0.052	-0.042 **	0.021
firm	4.442 ***	1.602	0.702	0.545	2.473 **	1.228
techtransfer industry	9.226 ***	3.321	1.242	0.869	2.192 ***	0.425
Log pseudolikelihood	-2039.088		-6269.135			
# censored observations	537		537			
ρ[equ. 1/2]	-		0.004 (0.019)			

Notes: Number of observations = 951. *** (**,*) indicate a significance level of 1% (5%, 10%). All models contain a constant. If we include unweighted publication and citation variables the signs and significance levels are similar.

Table 3.11: Estimation results from probit models with selection on “exit” (with inverse probability weights, only main variables presented)

	Model 1 (full sample)			**Model 2 (full sample)**			**Model 3 (age < 65)**			**Model 4 (age < 65)**		
	coef.	s.e.		coef.	s.e.		coef.	s.e.		coef.	s.e.	
Consulting activities												
consulting	0.023	0.010	**				0.029	0.012	**			
consulting^2	<-0.001	<0.001					-0.001	<0.001				
public consulting				0.005	0.016					0.016	0.025	
public consulting^2				0.001	<0.001					<0.001	0.001	
private consulting				0.023	0.012	*				0.027	0.011	**
private consulting^2				<-0.001	<0.001					<-0.001	<0.001	
# observations	951			951			909			909		
# consulting-active obs. (2nd stage)	414			414			392			392		
Log pseudolikelihood	-169,889			-169,475			-161,373			-160,992		
Wald test of indep. equations chi^2(1)	2.20			2.25			3.14*			2.67*		
α^{ρ}	0.771 (0.520)			0.570 (0.380)			0.954 (0.539)*			0.746 (0.456)*		

Notes: Number of observations = 951. Average marginal effects presented. *** (**,*) indicate a significance level of 1% (5%, 10%). All models contain a constant. If we include unweighted publication and citation variables the signs and significance levels are similar. Same set of controls variables as in Table 4 included, but omitted from this table.

Table 3.12: Comparison of different estimation methods (outcome equation without “exits”)

	OLS ln(pubs)	Tobit ln(pubs	Tobit # pubs	Poisson # pubs	Negbin # pubs	OLS ln(av.cits)	Tobit ln(av.cits	Tobit #
consulting	<-0.001	<-0.001	0.152	0.006	0.006	-0.016*	-0.016*	-0.359***
	0.008	0.008	0.178	0.009	0.009	0.008	0.010	0.135
consulting^2	<0.001	<0.001	-0.001	<-0.001	<-0.001	<0.001	<0.001	0.004**
	<0.001	<0.001	0.002	<0.001	<0.001	<0.001	<0.001	0.002
age	0.144***	0.144***	1.078	0.071	0.169 ***	0.067	0.074	1.023
	(0.046)	-0.045	-1.396	-0.056	-0.053	-0.046	-0.047	-0.732
age^2	-0.002***	-0.002***	-0.015	-0.001*	-0.002 ***	-0.001	-0.001	-0.010
	<0.001	<0.001	-0.013	-0.001	-0.001	<0.001	<0.001	-0.007
ln(average number of co-authors)	0.154	0.154	2.704	0.102	0.179	0.021	0.008	0.798
	-0.113	-0.109	-3.430	-0.100	-0.125	-0.120	-0.122	-1.892
ln(publications)$_{2002\text{-}2008}$ /	0.625***	0.625***	14.808***	0.705***	0.645 ***	0.401***	0.407***	5.183***
ln(av.citations)$_{2002\text{-}2008}$	-0.044	-0.043	-1.841	-0.048	-0.045	-0.052	-0.053	-0.792
no_publication_d / no_avcit_d	0.711***	0.711***	23.074***	0.681***	0.738 ***	0.512*	0.484*	8.248**
	-0.195	-0.189	-5.400	-0.254	-0.223	-0.264	-0.269	-3.758
ln(industry funding)$_{2002\text{-}2006}$	0.012	0.012	3.354	0.147	-0.197	0.687**	0.717**	13.234***
	-0.193	-0.187	-6.418	-0.219	-0.201	-0.294	-0.289	-4.295
ln(public funding)$_{2002\text{-}2006}$	0.160	0.160	2.049	-0.123	0.017	-0.075	-0.092	-4.039
	-0.144	-0.139	-3.380	-0.159	-0.161	-0.152	-0.153	-2.584
collaborative reach$_{2002\text{-}2008}$	0.047	0.047	-0.002	0.024	0.056	0.046	0.049	0.095
	-0.032	-0.031	-0.953	-0.037	-0.036	-0.036	-0.036	-0.528
ln(patents$_{pre2009}$)	0.102*	0.102*	3.227*	0.142***	0.178 ***	-0.047	-0.051	-1.688
	-0.059	-0.057	-1.672	-0.042	-0.057	-0.067	-0.069	-1.049
R^2/Log pseudolikelihood	0.68	-317.27	-1406.07	-1540.89	-1087.27	0.44	-368.08	-1299.26
sigma	-	0.62	15.69	-	-	-	0.71	12.69
# censored obs.	-	-	-	-	-	11	11	11

Notes: N = 337. Unlike in the main text of the paper all outcome variables are non-field-weighted here to show the comparison with count data model estimates (without selection stage). Coefficients presented. Robust standard errors below. *** (**,*) indicate a significance level of 1% (5%, 10%). All models contain a gender dummy, a dummy variable for co-authorship with industry, ln(peer group size), international visibility and rank dummies as well as a variable indicating university affiliation and a constant (not presented). *** (**,*) indicate a significance level of 1% (5%, 10%).

Robustness test for selection on unobservables

The identification strategy using selection equations as described above relied on a set of exclusion restrictions, which, while fulfilling the commonly applied statistical criteria, may be disputed for theoretical reasons. We therefore test the robustness of the findings presented above using an instrumental variable (IV) approach suggested by Lewbel (2012). This method does not rely on external IVs, but achieves identification through the generation of IVs based on heteroscedasticity. Despite its relative novelty, this approach has been used in several studies and the authors generally conclude that the results are comparable to those obtained from external instruments (see for instance, Emran and Hou 2013). In particular, in a two-stage model the potentially endogenous variables, the consulting shares in our case, are instrumented with variables generated such that they are uncorrelated with the product of heteroskedastic errors. We first test for the presence of heteroscedasticity in models for the overall consulting share as well as for public and private consulting separately. In all three cases, the Breusch-Pagan test strongly rejects the null of homoscedastic errors. The generated instrumental variables are relevant as indicated by the F-test of joint significance in the first stage and the models are not overidentified as indicated by the Hansen J test statistic for overidentification of all instruments. The results from the IV estimations are presented in Table 3.13 and they confirm the main effects from previous models.

Table 3.13: Instrumental variable (IV) models with heteroscedasticity-based instruments

	ln(field-weighted publications)	ln(field-weighted publications)	ln(field-weighted av.citations)	ln(field-weighted av.citations)
	coef. (s.e.)	coef. (s.e.)	coef. (s.e.)	coef. (s.e.)
consulting	-0.002		-0.003	
	(0.002)		(0.003)	
private consulting		-0.001		0.001
		(0.002)		(0.004)
public consulting		0.002		-0.005 **
		(0.003)		(0.003)
age	0.054 ***	0.054 ***	0.004	0.005
	(0.015)	(0.015)	(0.019)	(0.019)
age^2	-0.001 ***	-0.001 ***	(<-0.001)	(<-0.001)
	(<0.001)	(<0.001)	(<0.001)	(<0.001)
female	0.017	0.014	-0.014	-0.014
	(0.031)	(0.031)	(0.036)	(0.036)
ln(field-weighted $publications_{2002-2008}$	0.589 ***	0.587 ***	0.473 ***	0.472 ***
/ $av.citations_{2002-2008}$	(0.030)	(0.030)	(0.048)	(0.048)
no_publication_d / no_avcit_d	0.219 ***	0.217 ***	0.203 ***	0.196 ***
	(0.071)	(0.071)	(0.069)	(0.069)
ln(average number of co-authors)	0.090 **	0.091 **	0.061 **	0.061 **
	(0.040)	(0.040)	(0.028)	(0.028)
collaborative $reach_{2002-2008}$	0.031 ***	0.031 ***	0.017	0.018
	(0.010)	(0.010)	(0.012)	(0.012)
international visibility	0.094	0.099	-0.004	-0.001
	(0.091)	(0.090)	(0.100)	(0.100)
ln(industry $funding)_{2002-2006}$	-0.011	-0.018	0.151	0.116
	(0.096)	(0.095)	(0.121)	(0.120)
ln(public $funding)_{2002-2006}$	0.053	0.049	0.040	0.049
	(0.043)	(0.044)	(0.052)	(0.052)
ln(peer group size)	0.013	0.013	0.014	0.014
	(0.008)	(0.008)	(0.012)	(0.012)
university	-0.067 **	-0.058 *	-0.053	-0.053
	(0.030)	(0.031)	(0.036)	(0.036)
ln($patents_{pre2009}$)	0.055 ***	0.055 ***	-0.024	-0.024
	(0.020)	(0.021)	(0.019)	(0.019)
co-authorship industry	0.089 ***	0.088 ***	0.037	0.034
	(0.031)	(0.031)	(0.038)	(0.038)
# of observations	784	784	784	784
Underidentification test (Chi^2-test)	33.79 **	54.83 **	37.50 **	58.12 **
Sargan-Hansen J. (Chi^2-test)	14.75	31.65	23.52	38.56

Notes: Robust standard errors. *** (**,*) indicate a significance level of 1% (5%, 10%). All model stages contain rank and field dummies and a constant. First stage regression results are available upon request.

4. Research at the frontier of knowledge: The use of text similarity indicators for measuring scientific excellence

Abstract*

The identification of scientific excellence is of crucial interest to science and innovation for allocating scarce research resources to the most promising projects and individuals. Doing so, however, remains a challenge given the idiosyncratic nature and complexity of research outcomes and their evaluation. In addition to peer evaluations, decision makers often rely on citation-based indicators as they provide a quick and reliable way to study researchers' output. Citations however depend on many factors - they are highly field-specific, and due to their ex-post nature may not be suitable for the evaluation of recent research (MacRoberts and MacRoberts 1996). This study compares citations as a standard measure of scientific excellence to text-based similarity indicators by using natural language processing (NLP) techniques. The proposed similarity indicators are based on the idea that scientific proximity between individual scientists and verified knowledge frontiers can be traced through text-based similarity between scientific documents. We test this idea by using co-word analysis to calculate similarity scores for a sample of 1884 international scientists and two knowledge frontier definitions: academic prizes and European Research Council grants. Our comparisons suggest a high correlation between content-based similarity scores and citation-based indicators, and that both can be predicted by an individual's academic rank, institutional prestige and research funding. These correlations tend to be stronger for academic prize winners as proxies for knowledge frontiers than ERC funding recipients. We suggest that this may be due to the "deeper" and broader information base of this frontier measure. Further, text-similarity indicators might be an interesting alternative to identify excellence in younger scientists who exhibit lower citation counts simply due to their shorter career age and the cumulative nature of citations. Overall, the results suggest that text similarity approaches can be valuable to complement peer review and standard bibliometric indicators in future science evaluations.

Acknowledgements

The authors thank participants of the Max-Planck Institute for Innovation and Competition Brownbag Seminar and Henry Sauermann for helpful comments. This research was funded by the German Research Foundation (DFG), grant number HO 5390/1-1.

* This chapter is co-authored by Hanna Hottenrott.

4.1 Introduction

The identification of scientific excellence is of crucial interest to public science administrations that aim to allocate scarce research funds to the most promising projects and persons. Excellent scientists have high probabilities to contribute significantly to science by means of original ideas, findings, and pioneering work, thereby pushing back research frontiers and opening up new fields of knowledge (Tijssen et al. 2006).

Peer evaluation, as a quality control mechanism, is despite its costliness (Rowland 2002) or potential biases (Lee et al. 2013), still regarded as the most reliable way to identify excellent scientists (Chubin and Hackett 1990). Peer evaluations in science are usually augmented with bibliometric indicators. Such indicators provide a quick and reliable way to study the quantity (usually publications), the quality (usually citations) and outlets (usually academic journals) of scientific output by scientists and their governing institutions. To infer scientific excellence, scholars typically relate to the upper tail (e.g. top 10% or 1%) of the research quality distribution (Bornmann 2014, Bonaccorsi et al. 2017). Indicators of excellence include for example the number of highly cited articles, the number of articles in highly cited journals, or the number of highly cited authors employed by an organisation or located within a country (Vinkler 2007). Despite their usefulness, citation counts have been criticized to suffer, for example, from directed citations, cronyism, ceremonial citations, nepotism, negative citations, field specificities, and selectivity to cite others (MacRoberts and MacRoberts 1996; Meho 2007; Catalani et al. 2015).

To overcome disadvantages of peer review and citation indicators, new methodological approaches are needed. Especially content-based analysis (CBA) appears a promising route to achieve this. Publications provide a rich source of information which can inform us on the underlying research quality if we have the right tools and benchmarks for analysis and comparison. Despite recent advances in natural language processing (NLP), large scale computing infrastructures, and nearly abundant scientific documents "to mine" scientific communication, little is known about what evaluators can expect from publication contents for quality and excellence inference.

This study presents a comparative analysis of standard and new indicators of scientific excellence. In particular, this study addresses the research question of whether text-based similarity between publications of individual scientists in different scientific fields of (i.e. biology, chemistry, economics and engineering) and documents of validated knowledge frontiers can be used to evaluate scientific excellence. More specifically, we compare citation counts and text-based indicators, and study how they relate to alternative subject-specific measures of research quality, such as academic rank, institutional rank, and research budget. The comparison aims to validate whether text-based, citation-based and research related indicators provide a coherent picture, or if they point in perpendicular directions.

For the construction of indicators, we calculate document-document similarity scores between a sample of 1884 scientists and two knowledge frontier definitions. The first frontier definition is based on 575 recent science prize awardees and their scientific publications between 2011 and 2016. The second knowledge frontier is based on the project descriptions of 3114 prestigious research grants (European Research Council grants) awarded during the same time period. Both knowledge frontiers involve a highly competitive peer-review process, are based on recent achievements in advancing human knowledge, and thus appear suitable as a reference point of excellence in science. The underlying text data was obtained from the publication records of each sample author and each prize awardee. We merged each authors' titles, keywords and abstracts into one document per author (henceforth sample documents, frontier documents). For ERC projects, we downloaded the project information from the EU CORDIS database and combined title and project objective into one document per project. After this, we used common text mining techniques like filtering, tokenization, and term weighting to standardize the vocabulary for the comparison.

After pre-processing, we use co-word analysis to obtain similarity score between each sample document and each frontier document in the respective field using four binary (i.e. Jaccard, Dice, Russel, Simple Matching Coefficient) and four metric (i.e. eJaccard, eDice, Cosine, Correlation) similarity measures. Co-word analysis has been widely used in scientometrics for detecting themes of research (e.g. Law et al. 1988), paper recommender systems (Beel et al. 2016) and information retrieval in science (Doszkocs and Schoolman 1980; Bollacker 1998), similar researcher search (Gollapalli et al. 2012), or mapping of science (e.g. Small 1999, Sternitzke and Bergmann 2009, Schiebel 2012). This co-word analysis resulted in 16 average similarity scores per sample author, which are then compared to research quality related indicators.

The results show that the obtained average similarity scores are highly correlated with citations and other individual-level indicators of research quality. Higher similarity scores, namely authors with a higher proximity to the knowledge frontier, have higher research budgets, more senior academic ranks, and are employed at institutions with higher rankings. Models that use similarity scores based on academic prizes tend to have higher explanatory power than the models using ERC funding frontier. This may be due to the "deeper" and "broader" information base of this frontier indicator as it makes use of all publications of a price winner in the defined period of interest, rather than being project-specific as in the case of ERC grant descriptions. Furthermore, we find that the correlation between similarity scores and citation measures decreases with scientists' age. This indicates that especially for younger scientists the similarity scores might be an alternative quality indicator.

This study demonstrates the feasibility of text-based similarity scores for science evaluations that aim to identify scientific excellence. We present and validate a new method that allows to infer research quality and excellence from content-based analyses based on text similarity to frontier research documents. Within the scope of this investigation, we propose the utility of two knowledge frontiers for science evaluations that are academic prizes and ERC funding. We show that closeness to these knowledge frontiers highly correlates with citations which suggests that content-based analyses may indeed capture excellence to a similar extent. Although a variety of cases can be imagined where text similarity does not reflect a similar scientific alignment (or even quality), we argue that given the "right" reference points, pre-processing and parameters - text similarity approaches can be valuable to complement peer review and standard bibliometric indicators, especially for evaluating younger researchers for which citation analyses may be questionable.

4.2 Measuring scientific excellence

Scientific excellence plays a crucial role in research funding: "Managers of research institutions, funding agencies and (supra) national governments all face, for different reasons and goals, the same pervasive evaluative question: how can one define, recognize and compare 'science excellence' as objectively as possible?" (Tijssen et al. 2002, p. 382).

There are several challenges involved in the identification of excellent scientist. Unique expertise profiles of scientists, especially across disciplines, are difficult to compare "because disciplines shine under varying lights and because their members define quality in various ways." (Lamont 2009, p.2; Marx and Bornmann 2015). Further, growing specialization and the quick advance in many scientific fields make it difficult to judge scientific novelty and rigour from outside of the particular scientific domain. Therefore is peer evaluation as a qualitative method, despite its problems of costliness, (Rowland 2002) or potential biases (Lee et al. 2013), still regarded as the most feasible way to identify scientific excellence (e.g. Chubin and Hackett 1990). Also the highly individual character of discovery, the substantial amount of autonomy, multi-tasking, creativity, and immeasurable outputs are essential features of the academic environment, which make it difficult to monitor scientific progress closely, even for peers.

Peer evaluation is typically augmented by (quantitative) bibliometric indicators such as publication or citation counts, or journal impact factors. It is frequently stated that the research quality distribution based on citations is highly skewed, that 30% of papers never get cited (Tijssen et al. 2002, Bornmann 2014), and that the majority of influential scientific papers are authored by a relatively small number of (excellent) scientists (Narin, 1976, Seglen 1992). When measured by citation counts, scientific excellence usually refers to the upper tail of the research quality distribution, i.e. the top-1, top-5 or top-10 percentile (Bornmann 2014, Bonaccorsi et al. 2017, Tijssen and Kraemer-Mbula). However, citation counts are ex-post measures of excellence indicative for past achievements but not definitive for current or future excellence. And, "there exist many reasons for one research article citing the other, not all of which are directly related to the scientific quality of the cited work or the contributing researchers and institutions" (Tijssen et al. 2002, p. 383).

Bornmann (2014) reviews the many different bibliometric methods that were used to identify excellent papers over the last decades. He classifies bibliometric studies on research excellence into two basic approaches, using either an absolute or a relative threshold of citation counts, i.e. defining papers with more than X citations as excellent or defining papers in the X percentile as excellent. While the majority of the studies he analysed use a definition of excellence based on absolute numbers, he suggests that definitions based on a relative number should be preferred to allow for proper cross-field and cross-time-period comparisons. He adds that around half of the papers that used a quantitative method to identify the excellent papers, worked with percentile rank classes, e.g. the top 0.01; 0.1; 1.0 or 10.0% papers within the total set.

Early work by Narin (1976) suggests that bibliometric measures such as publications and citations correlate with non-literature measures. Likewise, bibliometric evaluations of papers, people, or institutions correlate well with peer evaluations (correlations between .6 and 0.8). Narin further discusses previous studies which draw correlations between bibliometric measures, and for example graduate school rankings (r=0.21), recognition (r=0.2-0.8), academic rank (r=0.56), department rankings (r=0.8-0.9), or peer rankings (0.93-0.98). Furthermore, bibliometric measures correlate well with independent measures of eminence (0.5-0.8), including awards, listings, academic rank or affiliation. He points to a study of Cole and Cole (1967) who show a correlation between the number of citations to the three most cited papers of a scientist and the number of awards (r=0.67), and prestige of the highest award (0.41), as measures of recognition.

Several authors propose to use multiple indicators in order to measure scientific excellence (e.g. Kostoff 1997; Martin 1996; Van Leeuwen et al. 2003; Vinkler 2007; Bonaccorsi et al. 2017). For senior Dutch academic researchers, Van Leeuwen et al. (2003) for example analyse received citations, average number of citations per publication, number of uncited articles, with international references such as the mean citation rate of the journal, mean field citations score, and a field-normalized journal impact indicator. They argue that each type of indicator reflects a particular dimension of research performance, stressing that single indicators may provide an incomplete picture of research performance. Also Bonaccorsi et al. (2017) suggest the use of multiple measures of excellence since different indicators capture different portions of underlying distributions. Examining differences in research excellence between European, Asian and US universities, the authors propose that "objective" and "subjective" indicators provide different perspectives. While universities might claim to be excellent in one or a few fields, more objective indicators that take into account international competition (such as the number of top 10% *worldwide* citations of a university) may give a totally different perception excellence. These studies suggest that a combination of the various types of indicators may provide evaluators with more valid and more useful assessment tools to estimate scientific excellence.

Tijssen et al. (2002) propose that *average* citation scores of academic Centres of Scientific Excellence appear to be inadequate to predict the production of highly cited (excellent) papers. Attention to the top 1% and top 10% most highly cited research papers is suggested to provide a more useful analytical framework. In their analysis of scientific excellence on the institutional level using Dutch bibliographic data, they conclude that highly cited papers do not necessarily equate to "breakthrough science or leading edge research with a pervasive widely spread impact" (p. 386), but that highly cited papers in the international research literature are indeed statistically valid proxy measures of scientific excellence. Moreover, the authors question how and to what extent citation measures actually capture the "intrinsic quality" of research and "scientific excellence" rather than international visibility across a range of scientific fields. They point to accolades such as prestigious science prizes and appointments from international committees as "the only truly accepted measure of scientific quality", due to a "lack of an adequately discriminating standard for conclusive validation studies" (p. 395).

Apart from publication and citation counts, also esteem indicators are used to inform quality assessments, for example in the UK research assessment exercises (Martin 1996). Such indicators include honours, awards and prizes; election to learned academies and academic professional associations; service to conferences; service to journals; and visiting fellowships, as an indication for excellence. Esteem indicators are non-bibliometric indicators of research quality which are based on the standing of an individual or of pieces of research within the academic community. However, their use as stand-alone metrics in the evaluation (and identification of excellence) was doubted by an Australian expert group since they are "markers of individual standing and of research-oriented workload, not of actual research quality" (Donovan and Butler 2007, 240). The authors further point out that such esteem indicators are important for disciplines where bibliometric indicators are difficult to apply, yet they have rarely been tested and their value for evaluation remains uncertain (Donovan and Butler 2007). Further, the authors measure esteem indicators by counting the sheer number of honours, awards and prizes. This approach however masks the relative importance of such awards by giving them equal weight, largely because of a lack of representative rankings of their importance. This fallacy has only recently been addressed by Zheng and Liu (2015) who survey the importance of international prizes among their recipients, relative to the Nobel Prize.

An alternative to using counts and rankings of awards and prizes, which we pursue in this study, is to identify awarded (or funded) scientists as a comparison group and then to use their publication records and project description content for science evaluations. This approach provides us with an external "reference point" or knowledge frontier, to which we can compare other scientists.

Frontier knowledge in science

Knowledge frontiers are discussed as part of the dynamics of knowledge generation and accumulation of economies. Saviotti (1998) for example states that newer parts of knowledge in a scientific discipline emerge at the frontier of knowledge as tacit knowledge (ideas and conjectures which are least codified), before they are gradually codified for communication in research publications, and later in text books (most codified). The author further states that each domain of knowledge (discipline, sub-discipline) has multiple frontiers which are explored by several researchers simultaneously, implying continuous shifts of the frontiers (Saviotti 1998). Continuous shifts further imply an increasing educational burden on successive cohorts of scientists, higher specialization and more teamwork to expand the frontiers at a constant rate (Jones 2009). Budreau et al. (2015) simply define knowledge frontiers by means of public scientific communication including publications, conferences, seminars, textbooks, graduate training, and other means that create a common stock of open knowledge. They term a commonly perceived knowledge frontier as "an envelope that demarcates what is currently known from what remains to be investigated" (p.2767). Persson (1994) distinguishes between the research front and the intellectual basis of research fields. He refers to cited articles as an intellectual base and citing articles as a research front, i.e., clusters of articles using similar parts of the intellectual base.

Scientific awards as indicators for excellence

Academic awards are an important source of motivation of scientists for outstanding academic performance (Lam 2011, Stephan 2011). Examples of academic awards range from best paper awards and honorary doctorates to the highly esteemed Nobel Prizes and invitations to national academies of sciences. Such awards can be of three general types: lifetime accomplishments over an entire career; current accomplishments in the most recent time period; and special accomplishments which are not covered by the other award types, e.g. field awards or awards for female scientists (Weisbrod and Hansen 1972). Depending on the awards' level of significance, they can have direct and indirect income effects for the scientist, but also shape the trajectory of their academics careers. While direct income effects usually stem from some sort of monetary gratification, e.g. prize money, indirect effects of academic awards add something to scientists' résumé that money can't buy: appreciation and recognition on the part of colleagues and the public (Frey and Neckermann 2010). Academic awards reveal to some extent talent, motivation and dedication, but also scientific excellence - characteristics which are typically hard to assess for outsiders (Frey and Neckermann 2008). Such indirect income effects help to build up a reputation, making the person's works known to a wider audience, and can facilitate access to external funding (Zuckerman 1992). In other words, they pave the way for cumulative advantages and Mathew's effect in science (Merton 1968).

However, according to Frey and Neckermann (2008), there seems to be "almost no serious empirical evidence on the effects of awards on (research) performance, mainly because the properties and effects of awards have rarely been studied by economists or by other social scientists." (p. 5). While some evidence on this relationship exists for corporate employee performance (e.g. Frey and Neckermann 2008), and for corporate invention incentives (Wright 1983), few studies investigate effects of academic awards, most notably Nobel Prizes, on scientific or inventive performance in public research.

Inhaber and Przednowek (1976) demonstrate the differences in scientists' perceived importance (measured by citations) of the work before and after the award of Nobel Prizes or invitations to the US academy of sciences. They find mixed patterns for these awards in measuring "the width of the 'halo' accorded them". While for Nobelists in medicine the citation rate decreases, rather than increases, after the receipt of an award or membership (the "halo" shrinks), recipients in physics and chemistry gained substantial increases in citations after the award years. New members of the US academy of science in medicine have a much greater increase in citation rate than the Nobelists, while the pattern is reversed for chemistry. They point to the distorting role of publicity or media attention on subsequent citation patterns, and an increased scientific visibility which has a strong correlation with the rate of citations. It follows that prizes, among other factors, can be seen a shock to the otherwise organic accumulation of citations, and that citation measures also capture components that are not directly related to the quality of the underlying scientific achievement, e.g. visibility, discipline or citation trends and dynamics, or even writing style.

Hirsch (2005) proposes the well-established h-index. The index reflects that an individual has authored at least *h* scientific documents with at least *h* citations, i.e. a combination of publication quantity and quality, to measure research quality. However only as an example, he states that Nobel Prize winners typically have substantial h-indices, thereby confirming prizes are useful reference point in science evaluations.

Charlton (2007a, b) argues that Nobel Prizes and plausible surrogates can be used as a scientometric measurement of "revolutionary science". He focuses on "revolutionary" biomedical science using the NLG metric (Nobel prizes, Lasker awards and Gairdner awards, 1992–2006) and the broader NFLT metric (Nobel prizes, fields medals, Lasker awards and Turing awards to identify centres of excellence and demonstrates a marked concentration of excellent scientists and institutions holding prizes in the USA. Iaria et al. (2018) analyse whether the collapse of scientific communication during World War I affected the number Nobel Prize nominations in Central countries[14] for the years 1905 to 1945. They propose to use Nobel Prize nominations as a new measure of research impact. The authors show that scientists in Central countries who relied on frontier research from abroad produced less Nobel Prize-nominated research, published fewer papers in top scientific journals, and introduced fewer novel scientific words in follow-on research. They use Nobel Prize nominations as an indicator for scientific breakthroughs, and argue that access to frontier knowledge significantly affects the production of basic science and, the development of new technology.

The use of scientific awards, however, also suffers from several shortcomings. For example, honours and prizes are not as frequently awarded, and thus do not provide a solid data basis for science evaluations, as compared to the abundance of publication and citation counts. Further, the selection of scientists as awardees is typically based on a peer-review which is usually not open for scrutiny. Many drawbacks of peer-review, e.g. nepotism or subjectivity, also apply for the selection of prize awardees by gate-keepers in the scientific community.

Research funding as an indicator for excellence

Likewise to academic awards, prestigious research grants can reflect excellence. Funding is essential for scientists' work, especially in the hard sciences, and contributes exceptionally to research outcomes (Stephan 1996; Hottenrott and Lawson 2017). Receiving financial support from an institution with high prestige, signals the ability to pass through a competitive peer-review process based on excellent research ideas (proposals) and vindication. In this study, we propose that being awarded with reputable research grants can also be assessed as a signal for scientific excellence. A fundamental difference of research funding to academic prizes, which are awarded in retrospect, is that research funding is awarded to scientists who signal to do great research in the future. However, the idea to use prestigious funding as a benchmark to identify scientific excellence has not been addressed in the previous studies.

Content-based analysis

Despite its infancy with respect to theory and methods for science evaluation, it seems that content-based analysis of scientific communication is a promising research avenue that allows evaluators to overcome certain (not all) shortcomings of peer evaluation and citations numbers. Content analysis of publications is, in contrast to peer evaluation, more scalable, cheaper, faster and less prone to evaluation bias, however it may only mechanically identify some of the complex, dynamic and often subtle patterns of research excellence. Citation counts on the other hand are also scalable, cheap, fast and objective, however they largely underlie journals as their gate keepers in their ability to reflect research quality. Further, they exclusively rely on external

[14] The study distinguishes Allied countries (among others United Kingdom, France, United States) and Central countries (Germany, Austria-Hungary, Ottoman Empire, Bulgaria).

validation mechanisms, while largely ignoring what has caused the citation and also what the achievement is about. Although not fully resolving the superficiality of citation measures, content analysis of scientific publications provides a new space to find patterns of relationships which are invisible for citation counts, and which might be useful in science evaluations. This depends on the ability of content analysts to distil meaningful patterns and relationships from the text to make them part of the research quality equation. One of such patterns lies in the text similarity between publications of individual scientists and documents of validated knowledge frontiers.

In this study, we explore the feasibility and plausibility of content-based indicators which rest on frontier knowledge for science evaluation. More specifically, we address the research question of whether text-based similarity between publications of individual scientists and publications of award or ERC grant winners can be used as an indicator for research excellence. To answer this question, we first describe the distributions of the obtained similarity indicators and their correlation with citation-based indicators. In a second step, we investigate the correlation of both citations and content-based similarity scores with independent indicators of research quality (i.e. research budgets, academic ranks and institution ranks). For doing so, we regress citations and obtained similarity scores on individual and institutional characteristics that are usually associated with closeness to the knowledge frontier.

4.3 Data

Our analysis relies on the expressiveness of titles, keywords, and abstracts of scientific publications and research project descriptions. These meta-information reflect each documents' crude content since, as other authors have pointed out, "the inherent nature of titles and abstracts is to describe the major contents of a paper" (Sternitzke and Bergmann 2009, p. 118).

Academic prize awardees

The first knowledge frontier definition is based on international academic prize awardees. We identify relevant academic prizes using Zheng and Lius' (2015) list of "important international academic awards"[15]. From this list, we take all available prizes in four focal disciplines to identify recent price recipients. In particular, we include 10 prizes in economics, business and finance; 34 prizes in life sciences (biology and biosciences and medicine); 11 in chemistry; and 54 in engineering to our study (for details, see Table 4.13 in the Appendix).[16] We then looked up the

[15] These awards are selected on three criteria: a) They honour individuals' contribution to the advancement of knowledge (i.e. research awards); b) that are not restricted "on nationality, and generally regardless of race, gender, age, religion, ethnicity, sexual orientation, disability, language, or political affiliation"; and are c) "generally granted by international organizations, national governments, renowned foundations, academic associations, national academies and learned societies".
[16] We exclude Nobel Prize winners, since they are typically awarded for life-time academic achievements rather than for recent academic achievements.

recipients' names for the five past award periods (usually annual or biannual recognition) and their respective Scopus identification numbers (Scopus ID).

After manual cleaning and disambiguation of names and affiliations, we downloaded all publication records which listed the researcher either as author from the Scopus database for the time period 2011-2016. If more than one Scopus ID for a given author was found, we merged their records into one document. We retain only peer-reviewed English language articles for the period 2011-2016. From the list of prize winners, we further remove the duplicate entries of those scientists who won more than one award in a discipline during this period. We also exclude those authors that did not have peer-reviewed articles in the focal time period and publications without either abstract or keywords. This selection resulted in 575 prize awardees of which 45% are active in engineering, 37% in biology and medicine, and 9% in each chemistry and economics or business. In the following, we refer to these authors as "prize frontier authors". To simplify the later comparisons, we combine all available titles, keywords and abstracts of a frontier author into single text documents. We refer to these documents as *frontier documents*$_{PRI}$.

ERC project descriptions

The second frontier definition is based on grants awarded by the European Research Council (ERC). The ERC is the most prestigious European funding organization with the aim to support long-term funding of curiosity-driven research at the frontiers of knowledge. The ERC is designed to support high-risk basic research and pioneering research without topical restriction. The selection of grantees is conducted by peer review panels composed of renowned scientists, with "scientific excellence" being the principal selection criterion.

We consider 3664 projects that were granted between 2011 and 2016 and which were tagged by at least one subject area.[17] This resulted in project descriptions for 1897 starting grants, 313 consolidator grants, 1430 advanced grants, and 24 synergy grants. We downloaded their meta-information from the EU Horizon 2020 framework website and merged the title of the project and its description into single text documents referred to as *frontier documents*$_{erc}$.

Sample authors

We use data of individual researchers collected through the "International Science Affiliations" project conducted at the Technical University Munich (TUM) in 2016. The sampled authors were randomly chosen from journals stratified by their eigenfactor score in four scientific disciplines: biology (27%), chemistry (31%), economics and business (20%), and engineering (23%); see appendix A for a detailed survey description. In the survey, the respondents were asked to answer research-related questions, especially about their employment situation, their institutional

[17] Some projects have multiple field tags what causes the reported sum in Table 4.2 to be larger.

affiliations and their resource sharing behaviour. The survey provides several control variables that profile the respondents, including country, age, gender, academic position, and research budget (Table 4.1). We exclude 23 individuals which appear as a principal investigator of the ERC project descriptions and two individuals which we identify to have won a prize from the sample.

The dataset contains survey responses from scientists in Germany (27%), Japan (30%), and UK (43%). We classified the respondents into four occupational ranks, i.e. junior scientists (4%), post-docs (26%), assistant professors (31%) and full professors (38%). The age of the respondents ranges between 25 and 88 years with an average age of 46 years. Among the respondents were 17% women. Using the provided annual research budget (with a median of 150.000€ and a mean of 4.7 million €), we create four budget categories, one for each quartile, with each quartile containing nearly 25% of respondents. In order to control for institutional quality, we further lookup the institution rank of each respondent using international und country-ranking based in the "Times Higher Education Rankings". We classified the host institutions into a three-tier system (Tier 1: 0.18%; Tier 2: 23%; Tier 3: 24%) plus one class for those institutions that were not ranked (36%).

We complemented the survey data with respondents' publication records by downloading their publication records listing the researcher as author from the Scopus database until 2016. The publications were restricted to English language articles in peer-reviewed academic journals. During the focal period 2011-2016, the respondents' publication number ranged between 1 and 237 with a mean of 19 and a median of 12. These documents received on average 187 citation ranging between a minimum of one and a maximum of 7332. The median number of citation in the sample is clearly lower with 77 citations indicating a heavily right-skewed citation distribution (skewness of 7.80). In the full sample, authors had on average eight citations per publication and a median of six. Further, we add the number of co-authors per publication to control for team size effects (Persson et al. 2004). In the following, we refer to these scientists as *sample authors*.

Table 4.1: Descriptive statistics

Variable	Unit	source	median	mean	s.d.	min.	max.
Research budget							
1st quartile	binary	Survey	0	0.25	0.43	0	1
2nd quartile	binary	Survey	0	0.25	0.43	0	1
3rd quartile	binary	Survey	0	0.26	0.44	0	1
4th quartile	binary	Survey	0	0.25	0.43	0	1
Academic rank							
junior	binary	Survey	0	0.04	0.20	0	1
postdoc	binary	Survey	0	0.26	0.44	0	1
assistant professor	binary	Survey	0	0.31	0.46	0	1
full professor	binary	Survey	0	0.38	0.49	0	1
Institution rank							
not ranked	binary	THE Ranking	0	0.36	0.48	0	1
tier 1	binary	THE Ranking	0	0.18	0.38	0	1
tier 2	binary	THE Ranking	0	0.23	0.42	0	1
tier 3	binary	THE Ranking	0	0.24	0.42	0	1
Controls							
age	count	Survey	45	46.21	10.85	25	88
female	binary	Survey	0	0.17	0.37	0	1
country							
Japan	binary	Survey	0	0.30	0.46	0	1
United Kingdom	binary	Survey	0	0.43	0.50	0	1
Germany	binary	Survey	0	0.27	0.45	0	1
scientific discipline							
biology	binary	Web of Science	0	0.27	0.44	0	1
chemistry	binary	Web of Science	0	0.31	0.46	0	1
economics	binary	Web of Science	0	0.20	0.40	0	1
engineering	binary	Web of Science	0	0.22	0.42	0	1
Publication information							
$\text{publications}_{2011\text{-}2016}$	count	Scopus	12	18.92	22.09	1	237
citations	count	Scopus	77	186.70	381.35	1	7332
citations per publication	fraction	Scopus	6	7.97	13.67	0	519
co-authors per publication	fraction	Scopus	5	5.42	2.98	1	45

Notes: Number of observations = 1884. Funding variables in million €, THE: Times Higher Education.

Collection statistics

We refer to collections as sets of documents from a specific data source and in a specific field. In Table 4.2, we provide an overview of the twelve collections used in this study. These collection consist of four sample and eight frontier document collections. For each collection, we provide the number of authors, articles, total citations, and three ratios.

The sample comprises of 1884 authors with 27% in biology and medicine; 31% in chemistry, 20% in economics and business, and 22% in engineering. These authors have written 25842 articles in total and received more than 112577 citations during the years 2011-2016. The number of articles per author and the number of citation clearly varies between fields. Chemists have for example three times more articles per author than scholars in economics and business (29 vs. 8.8). We further see that biologists and chemists receive on average nearly twice as many citations per article, and clearly more citations per author, than economists and engineers.

Regarding academic prizes, we see that especially in chemistry and economics, there are fewer international awards than for biology and engineering (see Zheng and Liu 2015 for the prize population). The total 575 academic prize awardees have produced 14516 publications between 2011-2016. These 575 distinguished authors received nearly as many total citations as the total 2128 sample authors (381855 vs. 395407). In Biology, Chemistry and Engineering, scientists at the frontier of knowledge reveal a higher scientific productivity in terms of articles per author, compared to the sample authors. Only frontier economists seem on average to publish less but are nonetheless awarded with a science prize (5.9). Also the scientific impact, measured by citations, is clearly higher for frontier authors. Prize-winning engineers, for example, received more than 8 times the number of citations per author (101 vs. 823), while for economists this ratio is still twice as high.

Table 4.2: Collection statistics

		Biology & Medicine	**Chemistry**	**Economics & Business**	**Engineering**	**Total**
Focal authors	count	502	576	383	423	1884
Number of articles	count	9145	16697	3381	6416	25842
Total citations	count	90975	193934	21602	45236	112577
Articles per author	ratio	18.2 (16.0)	29.0 (29.4)	8.8 (10.2)	15.2 (18.8)	18.9 (22.1)
Citations per article	ratio	10.0 (24.0)	9.6 (7.1)	5.1 (6.2)	5.9 (5.1)	8.0 (13.7)
Citations per author	ratio	181.2 (238.3)	336.7 (588.0)	56.4 (147.1)	106.9 (195.8)	186.7 (381.4)
Academic prize awardees	count	214	53	52	256	575
Number of articles	count	4655	1828	306	7727	14516
Total citations	count	121976	44057	5074	210748	381855
Articles per author	ratio	21.7 (25.7)	34.5 (38.5)	5.9 (5.4)	30.2 (41.2)	25.6 (37.9)
Citations per article	ratio	26.2 (57.4)	24.1 (38.7)	16.6 (27.1)	27.3 (57.5)	26.3 (47.0)
Citations per author	ratio	570.0 (792.4)	831.3 (1122.8)	97.6 (143.4)	823.2 (1834.4)	672.3 (1458.8)
ERC project descriptions	count	1166	1369	509	1345	3114

Notes: Publication records and project descriptions for the years 2011-2016, for the ratios we report the mean and standard deviation in parentheses.

The project descriptions of the four ERC collections – obviously – do not have bibliometric citations counts, which we can compare to the previous collections. However their magnitude reveals that economics and business projects are less often funded to be funded by European Research Council.

4.4 Method

The calculation of text similarity between *sample authors* and *frontier authors* is based on several natural language processing (NLP) techniques.[18] In the following, we use "term", and "word" interchangeably although a difference exists.[19]

We use co-word analysis to calculate the scientific proximity between sets of sample and frontier documents. Scientific proximity is a spatially visualized representation of how fields, subjects, publications and authors are related, based on ideational or cognitive proximity (in contrast to physical proximity, Small 1999).[20] Co-word analysis is a text mining technique that extracts words from documents, standardizes the vocabulary and builds a matrix of word co-occurrences between documents.

In scientific publications each field or subject has its own idiosyncratic language and uses different metaphors, technical terms and abbreviations. Therefore, it makes sense to only compare authors of the same (or similar) scientific domain to obtain reliable similarity scores. It follows that we map the frontier fields from the academic prizes and the ERC projects to the four sample disciplines, i.e. biology, chemistry, economics and business, and engineering (see Table 4.12 in the Appendix for details). Hereby we allow authors to only "connect" with documents of their own kind of vocabulary.

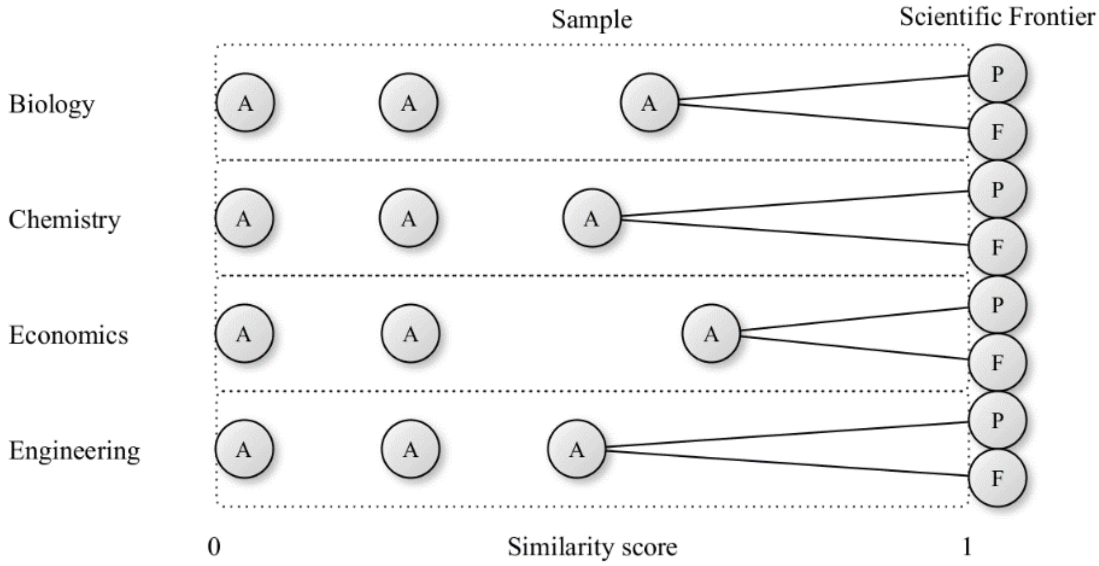
Figure 4.1 illustrates the calculation of normalized similarity scores. For each author (A) in the sample we calculate a score that is indicative of how "close" he/she is on average to all knowledge frontier documents, either prizes (P) or funding (F), in the respective scientific field. A high score means that a sample author has on average more words in common with all frontier authors and thus seems to be "closer" to the frontier than a sample author with a lower score. This builds on the assumption that two authors work on a similar topic, if they share a common vocabulary.

[18] We especially rely on the R packages tm and proxy.

[19] A term is a word that has meaning (semantics) and often refers to objects, ideas, or events whereas a word is only a component of language. Hence, all terms are words, but only some words can be terms.

[20] Despite scientific proximity, it exists a considerable body of literature that is concerned with technological proximity based on patent information (e.g. Bar and Leiponen 2012) or hybrids studies of scientific and technological proximity of technological (Magermann et al. 2010).

Figure 4.1: Document-document similarity between sample scientists and frontier science



Filtering and stemming

Rigorous pre-processing is crucial for co-word analysis and subsequent calculations. We start by removing punctuation, numbers and whitespaces from each document. Although numbers, especially in the life and natural sciences, can have discriminating value, they are without additional identifiers too generic for further use. For this reason we retain words only.

Next, we remove generic terms (stop words) from each document using three “stop word” lists. The first list is the SMART list from the R package tm. It includes 3589 terms such as “the”, “is”, “at”, “which”. Second, we construct a list of generic scientific vocabulary (like “results”, “method”, “implication” etc.) using the Academic Collocation List from Pearson Test of English Academic[21]. It comprises 2469 frequent lexical collocations in written academic English which are found in scientific jargon but not specific. A third list was constructed to remove unusual fragments that we encounter during the analysis such as non-English artefacts, publisher information or copyright statements.

We use the stemming algorithm by Porter (1980). Stemming methods refer to the simplification of morphological variants in terms, i.e. truncating words to their word stem (Frakes and Baeza-Yates 1992). For example, the word “categories” gets truncated to “categor”, in order to also match with “category”.

Document representation

One of the most widely used numerical representation of text documents is the vector space model (Manning and Schütze 1999). It represents each document as a high dimensional vector where

[21] https://pearsonpte.com/organizations/researchers/academic-collocation-list/

each dimension corresponds to a distinct term. A collection of m document vectors having a total of n terms will be represented by the $m \times n$ term-document-matrix A. Our goal is to transform A into a $m \times m$ similarity matrix S where $S_{i,j}$ gives some measure of the similarity between document vectors i and j. To do this, we specify four key parameters that condition the obtained similarity scores (see Table 4.3).

Table 4.3: Calculation parameter overview

Parameter		**Values**	**Description**
A	token size	unigram, bigram	term sequence (unigram, bigram, trigram, n-gram)
B	collection bounds	3	minimum collection frequency (absolute)
		.33	maximum collection frequency (in percent)
C	term length	3	minimum number of characters in a term
		inf.	maximum number of characters in a term
D	term weight[22]	augmented normalized tf	weight component for term frequency
		log(IDF)	weight component for inverse document frequency
		cosine normalization	weight component for document length

Token size

The first parameter refers to the unit of terms included in the matrix (tokenization). We can include every single term (unigram) or a fixed term sequence (bigram, trigram, n-gram) as a unit. A reinvestigation on the differences between unigram and bigram showed that the differences in the resulting similarity scores are rather small, i.e. result in a minor upward shift of the scores without substantially changing the score distributions. It follows that we primarily explore unigrams since, as Salton and Buckley (1988) point out, single term identifiers are preferable to more complex entities.[23]

Collection bounds and term length

In addition, we remove extremely frequent words as well as seldom words that lie within specific bounds of the collection frequency (Frakes and Baeza-Yates 1992). We chose a lower bound of 3 which means that we only retain terms that occur in at least three documents of the collection. This lower bound removes infrequent terms, misspelling and artefacts that do not help to distinguish one author from another. In the upper bound, we discard terms that occur in more than 33% of documents of the collection. From the term frequency distribution, we saw that most terms above this threshold do not characterize a scientific specialization of authors and therefore were

[22] Salton and Buckley (1988), "for technical vocabulary and meaningful terms (CRAN, MED collections), use enhanced frequency weights, first component n (augmented)"
[23] In Appendix B, we provide quantile-quantile plots on the differences between unigram and bigram tokenization for the obtained similarity scores. It appears that both seem to have a common distribution.

excluded. In another specification, we also try an upper bound of 10% to see how much this restriction affects the resulting scores. Although the differences were not large, we proceed with the more conservative approach of a 33% cut off. The third parameter is the term length. While we do not set an upper limit here, we require the terms to have at least three characters to be meaningful.

Term weight

Vector similarity functions depend on the choice of effective term frequency weighting schemes (Salton and Buckley 1988). Term frequency weighting is an important part of co-word analysis that intends to devalue non-discriminating terms while appreciating discriminating terms. We use Salton and Buckleys (1988) SMART weighting method where each term weight is the product of three components, i.e. term frequency, collection frequency, and vector length normalization. The resulting term weights, are then inserted into the cells of the term-document-matrices:

$$term\ weight_{t,d} = \mathrm{TF} \times \mathrm{IDF} \times \text{length normalization} \qquad (1)$$

Three considerations for are important for choosing an effective term weightings specification. First, Terms that occur often in a document appear to be important and thus should get higher weights. Second, terms that occur in many documents might not be discriminative and thus should get lower weights. Third, term frequencies in long documents should count less than term frequencies in short documents

The first component regards the frequency of term *t* in document *d.* It provides a good indication of the importance of a term. However, at a certain point, we're getting a diminishing return on its informative value since its relevance does not increase proportionally with frequency. This implies to scale down terms that appear too frequently. Instead of taking the raw term frequency count, we use the augmented term frequency for technical vocabulary and meaningful terms proposed by Salton and Buckleys (1988). It weights the $tf_{t,d}$ by the maximum term frequency, $max(tf_{t,d})$, in the document vector and gets standardized between 0.5 and 1. For example, if a term frequency is equal to the maximum frequency, the resulting weight equals 1. In contrast, if the term frequency is zero, the resulting weight equals 0.5.

The second component is the inverse document frequency (idf_t). It is the frequency of a term in a collection[24]. The term idf_t refers to a cross-document normalization that puts less weight on common terms, and more weight on rare terms. More specifically, the second part of the product divides the total number of documents in the collection (N) by the number of documents (n) to

[24] Salton and Buckleys (1988) use the expression "inverse collection frequency", which is more precise, however "inverse document frequency" (*idf*) is the standard label in the current literature.

which a term is assigned (Salton and Buckley 1988). We further take the log to deflate the effect of idf_t as in the first component.

The third component appears useful in systems with widely varying vector lengths. While some authors in this study have only one publication, others have hundreds. We correct such bias by equalizing the length of the document vector. This vector normalization prevents larger vectors from producing higher similarity scores just because they have more words that can potentially match between documents.

More specifically, we implement the following weighting systems, henceforth ATC for the augmented *tf/idf* with vector normalisation:

$$term\ weight_{t,d} = 0.5 + 0.5\ \frac{tf}{\max(tf)} \times \log_{10}(\frac{N}{\mathrm{n}}) \times\ \frac{1}{\sqrt{\sum tf^2}} \quad (2)$$

Calculating similarity scores

A large variety of measures have been proposed to express similarity, distance, or divergence between two statistical objects, e.g. tupel, vectors, sets or probability distributions (Lenz 2008; Deza and Deza 2009). These measures describe the statistical congruence between two document vectors that we wish to compare. The resulting similarity scores are usually high if two vectors have many common terms and low if not. While some coefficients are based on binary input, i.e. neglect the frequency with which a term occurs, others take into account the (weighted) frequency. The choice of similarity measure depends on the nature of data, the problem studied, and is not an exact science (Deza and Deza 2009).

Binary Similarity Models

Binary similarity measures do not use the term frequency directly and are rather based on four auxiliary variables (a-d). The binary models are defined by $t_{i,k} =\ 1$ if $tf_{i,k} > 0$ and $t_{i,k} = 0$ if $tf_{i,k} = 0$. Terms are thereof either present or absent in the document vectors. The auxiliary variables are defined as follows: For the i^{th} and j^{th} document, count $a_{i,j} =\ \sum_k t_{i,k} \times\ t_{j,k}$ is the number of mutual words present in both documents, $b_{i,j} =\ \sum_k t_{i,k} \times (1 - t_{j,k})$ and $c_{i,j} = \sum_k (1 - t_{i,k}) \times\ t_{j,k}$ represent words found in one document but not in the other. The number of terms that are mutually absent in both documents is denoted by $d_{i,j} =\ \sum_k (1 - t_{i,k}) \times (1 - t_{j,k})$ (see Table 4.4). Finally, n denotes the number of terms in the vectors. Using the auxiliary variables a-d, we implement the Jaccard Index, Sørensen–Dice Index, Simple Matching Coefficient, and Russel-Rao in our analysis (Table 4.5).

Table 4.4: Auxiliary variables for binary similarity models

	Term present in Document 1	Term absent in Document 1
Term present in Document 2	$a_{i,j}$	$b_{i,j}$
Term present in Document 2	$c_{i,j}$	$d_{i,j}$

Metric Similarity Models

We further include four metric similarity measures that use the term frequency, i.e. cosine similarity, extended Jaccard, extended Dice and Pearson correlation. Cosine similarity is, considered as the "state of the art" in similarity measurement. In metric similarity measures, usually the dot product $\sum_k x_{i,k}x_{j,k}$ of the term weight is used in the numerator while different variants of normalization are used in the denominator.

Table 4.5: Similarity measure overview

	Similarity measure	Description	Formula	
Binary	Jaccard index	simplest index, size of the intersection divided by the size of the union, ignores d	$Sim_{i,j}^{(Jacc)} =$	$\frac{a_{i,j}}{(a_{i,j} + b_{i,j} + c_{i,j})}$
	Sørensen–Dice index	similar to Jaccard, greater weight to shared terms $a_{i,j}$	$Sim_{i,j}^{(Dice)} =$	$\frac{2a_{i,j}}{(2a_{i,j} + b_{i,j} + c_{\ i,j})}$
	Russel-Rao	intersection divided by total number of terms	$Sim_{i,j}^{(Russ)} =$	$\frac{a_{i,j}}{n}$
	Simple Matching Coefficient	similar to Jaccard index, takes terms into account that occur in none of the two documents	$Sim_{i,j}^{(SMC)} =$	$\frac{a_{i,j} + d_{i,j}}{n}$
Metric	Cosine similarity	state of the art, computes similarity as the vector normalized dot product of X and Y	$Sim_{i,j}^{(Cosi)} =$	$\frac{\sum_k x_{i,k}x_{j,k}}{(\sum_k x_{i,k}^2 \sum_k x_{j,k}^2)^{1/2}}$
	Extended Jaccard index	extension of the Jaccard index to metric data, equivalent to the binary version when the term vector entries are binary	$Sim_{i,j}^{(eJacc)} =$	$\frac{\sum_k x_{i,k}x_{j,k}}{(\sum_k x_{i,k}^2 + \sum_k x_{j,k}^2 - \sum_k x_{i,k}x_{j,k})}$
	Extended Sørensen–Dice index	extension of the Sørensen-Dice index to metric data	$Sim_{i,j}^{(eDice)} =$	$\frac{2\sum_k x_{i,k}x_{j,k}}{(\sum_k x_{i,k}^2 \sum_k x_{j,k}^2)}$
	Pearson Correlation	formally identical to the cosine similarity, invariant to shifts	$Sim_{i,j}^{(Corr)} =$	$\frac{\sum_k x_{i,k}x_{j,k}}{(\sum_k x_{i,k}^2 \sum_k x_{j,k}^2)^{\ 1/2}}$ for centered weights

Procedure

With the pre-processed document vectors and the parameters A-D, we create a term-document matrix for each sample collection in a specific field and the respective frontier document collections (Table 4.6).

Table 4.6: Collections for similarity calculation

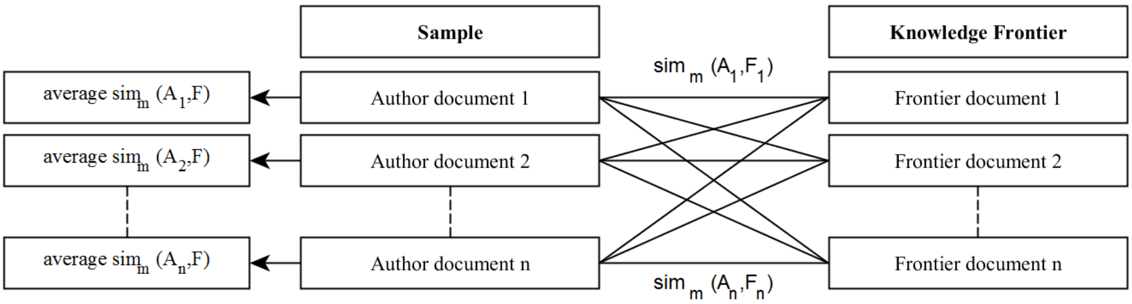
	Academic prize collections	**ERC collections**
Focal author collections	$sample_{bio}$ vs. $prize_{bio}$	$sample_{bio}$ vs. $funding_{bio}$
	$sample_{che}$ vs. $prize_{che}$	$sample_{che}$ vs. $funding_{che}$
	$sample_{eco}$ vs. $prize_{eco}$	$sample_{eco}$ vs. $funding_{eco}$
	$sample_{eng}$ vs. $prize_{eng}$	$sample_{eng}$ vs. $funding_{eng}$

Based on these matrices, we calculate similarity scores between each sample document *i* and each frontier document *j*, for each knowledge frontier *f*, using eight similarity measure *m* in each scientific domain *d* as follows:

$$Avg.Similarity\ Score_{i,f,m} = \frac{1}{N_j} \sum_{j=1}^{J} sim_m(Doc_{i,d}, Doc_{j,d,f}) \qquad (3)$$

Figure 4.2 illustrates the calculation. For each author document we calculate the pairwise similarity to all *frontier documents*$_{pri}$ and *frontier documents*$_{erc}$. From this calculation, we obtain 16 average similarity scores for each of the 2128 sample author (8 measures and two frontiers). To make the scores comparable across different measures, we normalize them by setting the highest resulting similarity score to one, the lowest score to zero, and all other scores relative to them (min-max normalization).

Figure 4.2: Similarity calculation procedure



Empirical model

In the following analysis, we use the obtained average similarity indicators in four basic OLS regression models. These models are identical with respect to the dependent variables and only vary by the dependent variable. We test whether a) *Avg. Similarity Score*$_{pri}$ b) *Avg. Similarity Score*$_{erc}$, c) *ln(citations*$_{total}$*) or ln(citations*$_{per\ article}$*)* are explained by the same characteristics typically found in excellent scientists. As independent variables, we use scientists', academic rank, institution rank, and research budget. Such ordinal variables are limited in their interpretation because they provide overall tendency without revealing intra-group differences. Therefore we split each independent variable into categories to allow for a more nuanced description. Accordingly, we construct a category for each quartile of the research budget (*1st – 4th quartile*), for each academic rank (*junior, post-doc, assistant professor and full professor*), and for each institution rank (*tier1, tier 2, tier 3, not ranked*). We then add these as dummies to the regression model. More formally, these models are constructed as follows:

$$research\ quality_i = \beta_0 + \beta_1 research\ budget_i + \beta_2 academic\ rank_i + \beta_3 institution\ rank + \sum_{n=4}^{k} \beta_n\ controls_i + u_i \qquad (4)$$

A set of control variables are included that have been shown to affect publication outcomes, such as age and gender, country, field (Toole and Czarnitzki, 2010; Mairesse and Pezzoni, 2015). Moreover, we add the number of co-authors per paper as a control to the regression models, since scientists with many co-authors obtain more citations than scientists who publish with fewer co-authors (Persson et al. 2004).

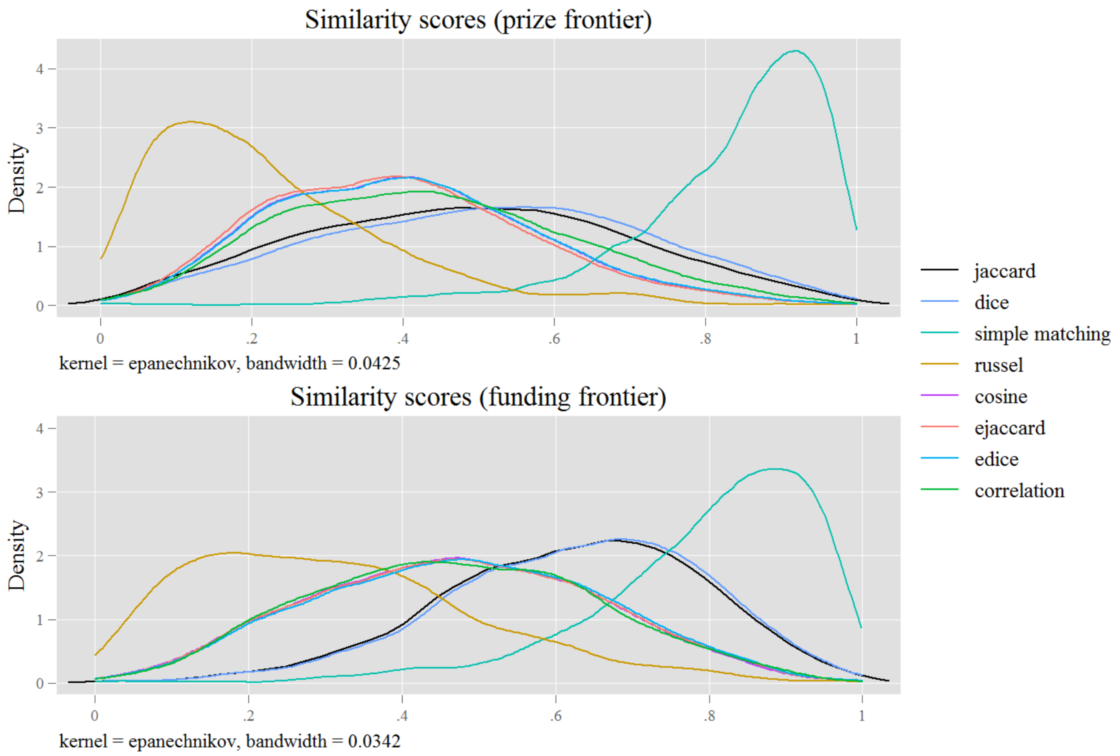
4.5 Average similarity scores

In this section, we describe the average similarity scores based on one specific parameter setting, i.e. augmented weighting with cosine normalization; unigram as token size; 3 characters as the minimum term length; a minimum term frequency of 3 and a maximum term frequency of 33%. An overview of the normalized distribution for each similarity score gives Figure 4.3, where we plot the kernel density of eight similarity measures and both knowledge frontiers. Kernel density estimation is a non-parametric method to estimate the probability density function of random variables, in this case similarity scores.

In the case of the prize frontier (top of Figure 4.3), it becomes apparent that most average similarity scores follow a symmetric and flat normal distribution. This normality is confirmed by a visual test for normality using quantile-quantile-plots (see Figure 4.6 in the Appendix). However there are two exceptions. The scores which are based on the Russel index are right-skewed (brown line), and those of the simple matching coefficient are left-skewed (green line), while both are more concentrated towards the lower and higher end of the score distribution. A deeper dive into their formulas in Table 4.5 reveals that this deviation is the result of incorporating "n", i.e. the total number of terms in the underlying highly sparse term-document matrix. Both measures, in a different way, make use of the total number of terms which results in similarity scores which show more extreme values. In contrast, the other two binary and four metric measures, are more centred on 0.5 (see Table 4.11 in the appendix for details).

A similar picture emerges for similarity score distributions using the funding frontier, but also for alternative parameter specifications.[25] While the density of the Russel index appears more normally distributed and less steep than in the case of academic prizes, the simple matching coefficient does not deviate much (lower part of Figure 4.3). The other two distributions based on binary measures (Jaccard, black line and Dice, blue line) peak in the third quantile and appear to be left-skewed. The four similarity scores based on metric measures do not differ much from those in academic prizes, however they are more symmetric and peak around the mode.

Figure 4.3: Estimated distributions of normalized average similarity scores (N=1884)



For both knowledge frontiers, it is noticeable that the binary measures Dice and the Jaccard, and their metric counterparts, extended Jaccard and extended Dice, follow each a similar distribution. The cosine similarity, which is a widely used similarity measure in text-based approaches, is perfectly hidden by the eDice measure, indicating that they capture the same information. The Pearson correlation measure (which is based on the cosine similarity for centred weights) also fits into this coherent group of measures. With other words, we might expect different effects in the following analyses for the deviating first group (Simple matching and Russel index), and the second more or less homogeneous group of similarity measures (Jaccard, Dice, eDice, eJaccard, cosine and correlation).

---

[25] To see how much these distributions depend on the parameter settings, we vary the token size between unigram and bigram, and the maximum term frequency between 33% (less restrictive) and 10% (more restrictive). See Figure 4.7-4.9 in the Appendix for details.

Some of the differences between prize and funding frontier might be explained by the data generating process. For the funding frontier, we obtain project descriptions which are short (usually the length of an abstract) but more numerous (cover a broader scientific spectrum). Average similarity score to the academic prize frontier, in contrast, are based on fewer comparisons (less prize winners), but with more profound contents (all titles, keywords, and abstracts of one author are merged into one document).

4.6 Results

Correlation analysis

This section describes the relationship between the obtained similarity scores and a set of commonly found indicators of research quality. More precisely, we compare the average similarity scores with scientists' citation counts, academic rank, institution rank, and annual funding budget. The basic idea of this comparison is to see whether the scores actually correlate with what has previously been related to research quality or even excellence.

Figure 4.4 provides several scatter plots which display similarity scores and citations per article as logged variables. These plots are separated by eight similarity scores and both knowledge frontiers. To better understand these 1884 data points, we add two fitted lines (linear and quadratic), a beta coefficient, and the R^2 from a univariate linear regression. Most plots show a positive relationship between similarity scores and citations per article. Only the simple matching coefficient has a negative correlation and most data points are clustered in the upper score range. Also the data points of the Russel index cluster in the lower range). This is likely to be caused by their deviant distribution which we describe in the previous section. For the academic prize frontier, the explained variance R^2 ranges between 8% (correlation) and 15% (Jaccard, Dice, Cosine, eJaccard and eDice). The explained variance of the score using the funding frontier ranges between 8% (Jaccard and Dice) and 21% (Russel). From the R^2, it appears that binary indicators based on the prize frontier explain citations per articles better than metric indicators. For the funding frontier, the opposite can be observed, i.e. the metric indicators have a higher explanatory power, with only the Russel index as an exception.

We repeat these scatter plots with the log of total citations in Figure 4.5. Since both citation measures are based on the same information, we find an identical positive relationship between scores and citations. However, the explained variance is clearly higher for total citations. They range between 26% (correlation) and 50% (Jaccard, Dice) when using the prize frontier as a reference. For the funding frontier a different picture emerges. While most measures have a positive correlation, their explained variance ranges between 2% (Jaccard) and 62% (Russel). Especially the binary measures Jaccard and Dice appear to have only a weak relationship with the log of citations. When using metric similarity measures, most of them have at least a moderate correlation with citation counts.

Overall we observe three things from the scatter plots. First, there seems to be a much higher correlation between the scores and total citations, rather than citations per article. Second, Metric similarity measures seem to provide a more homogeneous picture of this relationship while the binary ones are less consistent (especially simple matching and Russel). Third, the relationships between similarity scores and both citation measures seems to be better explained by the R^2 measure when using the prize frontier rather than the funding frontier.

Figure 4.4: Correlation between similarity scores and citations per article

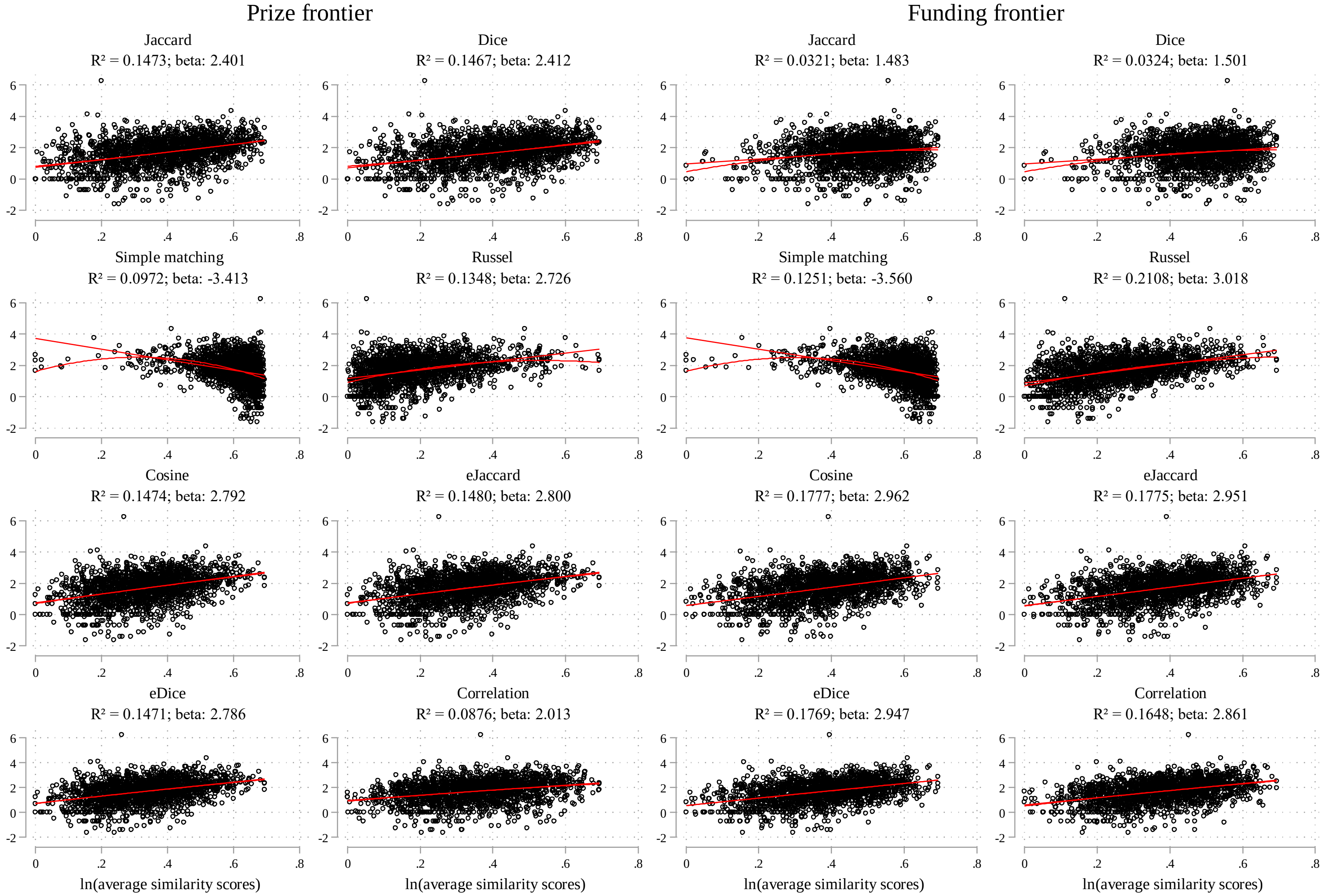
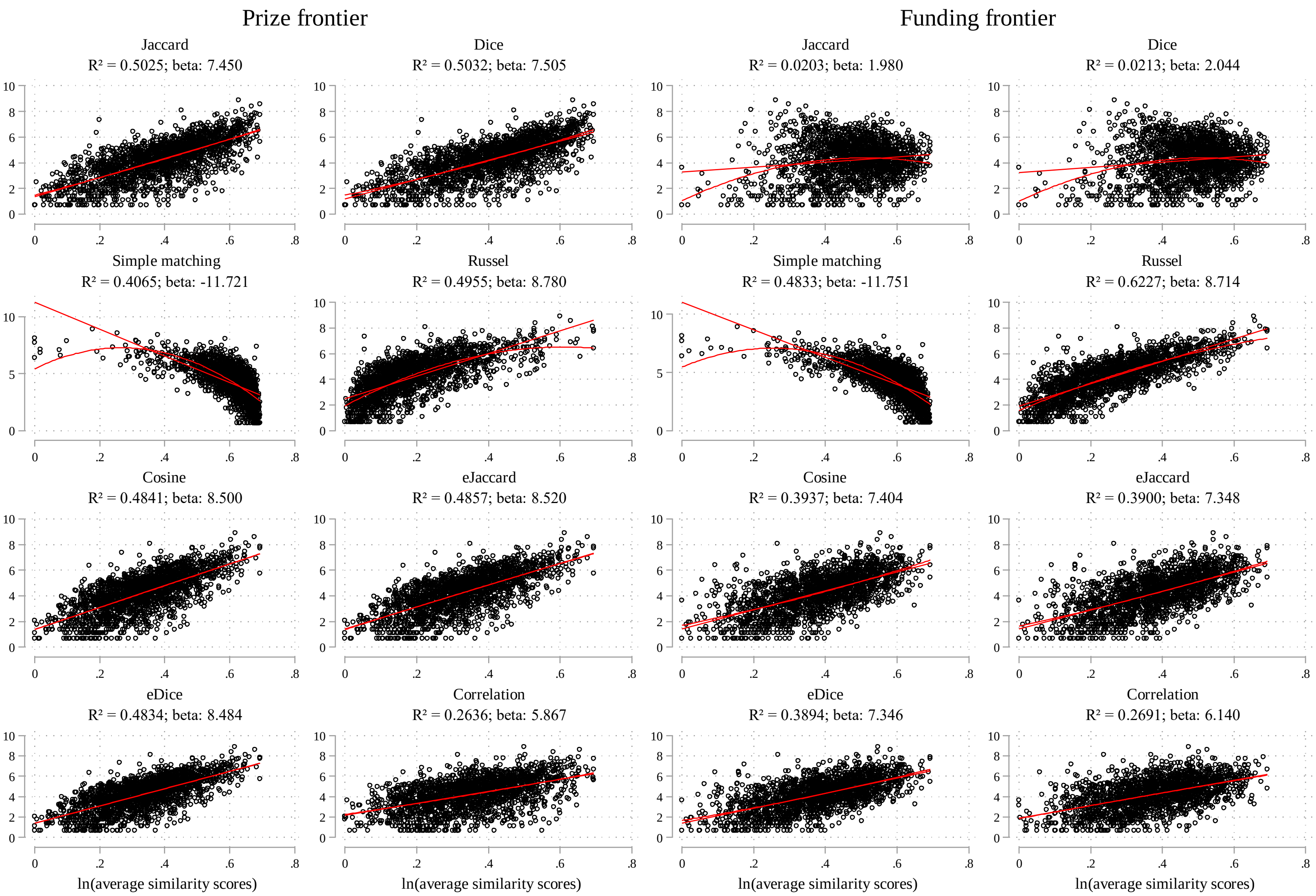


Figure 4.5: Correlation between similarity scores and citations



Decreasing correlation with higher age

Younger scientists, such as doctoral and postdoctoral students, have a structural disadvantage when their research quality is gauged by citation indicators. This is because citations largely depend on scientific visibility, which they typically lack. They have less papers out in the field, they get less collaboration requests, they have lower research budgets to visit conferences or to organize symposia, and are therefore, compared to established scholars, less visible in their scientific community (c.f. the cumulative advantage, Merton 1973). Comparing newcomers with established scientists by citation counts appears to be inequitable, if citation measures are not adjusted / weighted according to the individual career age (which is usually rarely done). For groups of scientists where citation indicators are less compelling, such as newcomers or in scholars in social sciences, arts and humanities, content-based indicators might be a substitute.

To test this idea for the obtained similarity scores, we analyse the interaction effect of age and the similarity score with respect to citations. From Tables 4.7 and 4.8, we find that the interaction effect between score and age is negative and significant for all metric similarity scores. For the binary similarity measures, the Jaccard and Dice index are insignificant while the simple matching and Russel index are strongly significant (not shown here). This means that the correlation between text-similarity score and citation indicators decreases for older scholars. In other words, the correlation between similarity scores and citations is stronger for young scientists and decreases with age.

This suggests that text similarity to frontier knowledge can be a valuable substitute when citation counts have limited expressiveness.

Table 4.7: The moderating effect of age (prize frontier)

	ln(citations)			
	cosine	**ejaccard**	**edice**	**correlation**
similarity score	7.078*** (.594)	7.164*** (.598)	7.048*** (.594)	5.193*** (.683)
age	.022*** (.005)	.022*** (.005)	.022*** (.005)	.035*** (.006)
similarity score ## age	-.026** (.012)	-.027** (.012)	-.026** (.012)	-.026* (.014)
_cons	.824*** (.254)	.863*** (.250)	.828*** (.255)	.866*** (.303)
observations	1884	1884	1884	1884
R^2	.49	.49	.49	.29

Notes: *** (**,*) indicate a significance level of 1% (5%, 10%).

Table 4.8: The moderating effect of age (funding frontier)

	ln(citations)			
	cosine	**ejaccard**	**edice**	**correlation**
similarity measure	7.299*** (.621)	7.250*** (.616)	7.230*** (.616)	6.807*** (.651)
age	.045*** (.007)	.046*** (.007)	.046*** (.007)	.056*** (.007)
similarity measure ## age	-.049*** (.013)	-.049*** (.013)	-.049*** (.013)	-.056*** (.014)
_cons	-.237 (.317)	-.265 (.316)	-.292 (.319)	-.363 (.336)
observations	1884	1884	1884	1884
R^2	.42	.42	.42	.32

Notes: *** (**,*) indicate a significance level of 1% (5%, 10%).

Regression analysis

Our main goal is to validate the use of text-based indicators for science evaluations. In this section we test whether the proposed similarity scores correlate with other individual-specific research quality indicators. In Tables 4.9 and 4.10, we provide OLS regression models with ten different dependent variables and the same set of independent and control variables. The first two models seek to explain factors that empirically affect scientific impact, quality or excellence with the log of citations and the of log citations per article as a proxy for these quality indicators. In models three to eight, we use average similarity scores as the dependent variable, one model for each similarity measure respectively. Similar to previous descriptions of the scores and the correlations, the simple matching and the Russel index seem to take a special role and their interpretation needs some caution. Their distributions, correlations and regression coefficients differ from other established measures such as Jaccard, Cosine and Correlation. Since the simple matching coefficient is always negative when other measures are positive, with a similar magnitude, we interpret its absolute values rather than its algebraic sign.

In Table 4.9, we compare citation models to similarity score models for the prize frontier. Scientists with a relatively high research budget (in the 4th quartile) have on average more total citations and more citations per article, than scientists with a very low research budget (1st quantile). In contrast, scientists with low (2nd quantile) and high (3rd quantile) research budgets do have clearly more citations than scientists with very low research budgets, although their regression coefficients are positive. For the similarity score models, we observe that scientists with high and very high research budgets (3rd and 4th quartile) have higher similarity score compared to scientists with very low research budgets. As a result, we find that higher research budgets show a positive correlation with citations and text-based similarity indicators.

We also observe a strong relationship between academic ranks and the considered outcome variables. Full professors have on average more citations and more citations per article than their junior colleagues without PhD. Also medium ranked scientists like post-docs and assistant professors exhibit clearly more citations and citations per article than junior researchers. A similar perspective emerges from the similarity score models. Full professors have on average the highest scores, compared to junior researcher. Scientists in postdoctoral positions and assistant professors, also show a positive effect on similarity scores. We find that citation indicators and similarity scores increase with scientists' academic ranks.

A coherent picture is also visible for the institution ranks. Scientists from high and medium ranked institutions (Tier 1 and Tier 2) exhibit higher citation indicators and similarity scores compared to scientists of unranked institutions. Tier 1 scientists have more total citations and more citations per article than scientists from unranked institutions. The similarity scores also correlate positively with institutional ranks.

In Table 4.10, we provide the same empirical setting as in Table 4.9, however with similarity scores obtained for the funding frontier. The first two models are equal to those in Figure 7 and only repeated for comparison. Here, we find that the similarity scores are also positively related to research budgets, academic rank, and institution rank. However the magnitude of the quality indicators varies slightly.

It further appears that the model which explains citations per article has a higher explanatory power (R^2) than the model for total citations. Further, the R^2 varies widely for similarity scores based on different measures, ranging between 0.37 (correlation) and 0.76 (simple matching) in the case of the prize frontier. The total variance explained by the scores based on the funding frontier are clearly lower, than those of the prize frontier. They range between 0.22 (Jaccard and dice) and 0.76 for simple matching.

From the regression models in Table 4.9 and 4.10, we can infer two broader finding. First, all three independent variables are found to be positively related to the average similarity scores. It turns out that higher budgets and higher ranks are associated with higher similarity scores. This is similar to what we expected and found for citations as a measure of scientific excellence. It seems that citations and similarity scores tend to measure a similar phenomenon, which we assume to be scientific excellence. Second, we find that the models which use the prize frontier altogether have a higher explanatory power. This suggests that “deeper” (more titles, keywords, and abstracts of one author merged into one documents) and “broader” (covering a spectrum of specializations) knowledge frontier measures would be favourable when using similarity scores to identify scientific excellence.

Control variables

The models incorporate several control variables that have been shown to affect publication outcomes. The number of published documents has a positive effect on all citation and similarity indicators. This effect is however small in contrast to previously described independent variables. We further see that additional co-authors per publication increase the citation and similarity indicators (as was shown by Persson et al. 2004). Also age has a positive effect on citations per article and a slight effect on the similarity scores. However, the age squared is negative and significant indicating a decreasing effect for older scientists. While we do not find a significant effect of gender on the citation indicators, we find that women have on average lower similarity to the knowledge frontiers. Finally, we observe some significant differences for the four disciplines and three countries. It appears that UK and German scientists on average attract more citations and citations per article, than Japanese scientists. Regarding scientific discipline, we observe that scientists in economics attract clearly less citations than scientists in engineering, while biologists and chemists are more often cited on average than engineers. With respect to the similarity scores, the discipline indicators provide a mixed and somewhat inconsistent pattern.

While some of the binary indicators are strong and significant in economics (compared to engineering), this impression is reversed for biology with metric indicators showing strong relationships and binary scores being small and insignificant. Chemists, on the other hand show consistently lower similarity scores than engineers. This inconsistency in the coefficients of disciplines however disappears when the funding frontier is used, rather than the prize frontier.

Overall, we find that all three independent quality indicators (research budget, academic rank, and institution rank) show a positive correlation with both sets of citation and similarity indicators. With respect to the frontier definition, we find a remarkable resemblance between the coefficients of the prize and funding frontier. Their coefficients deviate only slightly, tend to be lower for the funding frontier but do not contradict each other in terms of algebraic sign. Further, we find that the explanatory power (R^2) of the models using the prize frontier are higher than those of the funding frontier. Finally, we see the previously stated differences of the similarity measures confirmed. The simple matching and the Russel index do not contradict the other measures, however they tend to provide somewhat the extreme values in the spectrum of regression coefficients and R^2.

Table 4.9: OLS regression (prize frontier)

	ln(citations)	ln(citations per article)	jaccard	dice	simple matching	russel	cosine	ejaccard	edice	correlation
research budget					1st quartile = reference category					
2nd quartile	.042 (.055)	.106 (.068)	.014 (.010)	.015 (.010)	-.001 (.005)	.004 (.006)	.009 (.008)	.009 (.008)	.009 (.008)	.013 (.011)
3rd quartile	-.003 (.057)	.048 (.071)	.028*** (.011)	.028** (.011)	-.003 (.005)	.016** (.006)	.028*** (.009)	.029*** (.009)	.029*** (.009)	.047*** (.012)
4th quartile	.137* (.072)	.346*** (.089)	.070*** (.013)	.072*** (.013)	-.013** (.006)	.027*** (.007)	.046*** (.011)	.045*** (.011)	.046*** (.011)	.052*** (.014)
academic rank					junior = reference category					
postdoctoral position	.279** (.124)	.827*** (.141)	.131*** (.018)	.137*** (.019)	-.034*** (.007)	.053*** (.010)	.098*** (.014)	.097*** (.014)	.099*** (.014)	.109*** (.016)
assistant professor	.317** (.129)	.972*** (.148)	.159*** (.020)	.166*** (.020)	-.038*** (.008)	.062*** (.010)	.118*** (.015)	.116*** (.015)	.119*** (.015)	.135*** (.018)
full professor	.411*** (.134)	1.175*** (.161)	.196*** (.022)	.202*** (.023)	-.055*** (.009)	.090*** (.012)	.150*** (.017)	.148*** (.017)	.151*** (.017)	.165*** (.020)
institution rank					not ranked = reference category					
tier 3	-.010 (.053)	.072 (.067)	.013 (.010)	.013 (.010)	-.003 (.004)	.002 (.006)	.003 (.008)	.003 (.008)	.003 (.008)	-.007 (.010)
tier 2	.094** (.046)	.195*** (.057)	.032*** (.009)	.033*** (.009)	-.009** (.004)	.012** (.005)	.020*** (.007)	.020*** (.007)	.020*** (.007)	.018* (.010)
tier 1	.205*** (.048)	.294*** (.060)	.036*** (.009)	.038*** (.009)	-.008* (.004)	.014** (.006)	.033*** (.007)	.032*** (.007)	.033*** (.007)	.043*** (.010)
publications	.007*** (.001)	.035*** (.003)	.005*** (.000)	.005*** (.000)	-.005*** (.000)	.006*** (.000)	.005*** (.000)	.005*** (.000)	.005*** (.000)	.004*** (.000)
co-authors per article	.059*** (.009)	.090*** (.012)	.014*** (.002)	.014*** (.002)	-.007*** (.001)	.009*** (.001)	.011*** (.002)	.011*** (.002)	.011*** (.002)	.011*** (.002)
age	.010 (.014)	.037** (.017)	.008*** (.003)	.009*** (.003)	-.004*** (.001)	.004*** (.001)	.006*** (.002)	.005*** (.002)	.006*** (.002)	.003 (.003)
age²	-.000 (.000)	-.000** (.000)	-.000*** (.000)	-.000*** (.000)	.000*** (.000)	-.000*** (.000)	-.000*** (.000)	-.000*** (.000)	-.000*** (.000)	-.000** (.000)
female	.052 (.052)	-.055 (.061)	-.019** (.009)	-.019** (.010)	.000 (.004)	-.008 (.005)	-.016** (.007)	-.016** (.007)	-.016** (.007)	-.025** (.010)
country					Japan = reference category					
Germany	.536*** (.062)	.705*** (.080)	.145*** (.012)	.147*** (.012)	-.049*** (.006)	.080*** (.007)	.118*** (.010)	.116*** (.009)	.118*** (.010)	.133*** (.013)
United Kingdom	.634*** (.062)	.749*** (.077)	.126*** (.012)	.127*** (.012)	-.043*** (.005)	.070*** (.007)	.102*** (.009)	.100*** (.009)	.102*** (.009)	.110*** (.013)
field					engineering = reference category					
biology	.393*** (.054)	.406*** (.067)	.004 (.010)	-.009 (.010)	-.005 (.005)	-.010* (.006)	-.047*** (.008)	-.044*** (.008)	-.045*** (.008)	-.115*** (.010)
chemistry	.429*** (.050)	.587*** (.065)	-.054*** (.009)	-.059*** (.010)	.078*** (.004)	-.069*** (.005)	-.090*** (.007)	-.085*** (.007)	-.090*** (.007)	-.121*** (.010)
economics	-.270*** (.066)	-.522*** (.084)	.086*** (.014)	.073*** (.014)	-.039*** (.007)	.065*** (.008)	.009 (.011)	.014 (.011)	.011 (.011)	-0.000364
_cons	.132 (.335)	.454 (.398)	-.169*** (.058)	-.154** (.060)	1.125*** (.024)	-.145*** (.031)	-.046 (.047)	-.053 (.046)	-.047 (.047)	.080 (.061)
observations	1884	1884	1884	1884	1884	1884	1884	1884	1884	1884
R²	.31	.63	.56	.55	.76	.72	.58	.58	0.58	0.37

Notes: *** (**,*) indicate a significance level of 1% (5%, 10%).

Table 4.10: OLS regression (funding frontier)

	ln(citations)	**ln(citations per article)**	**jaccard**	**dice**	**simple matching**	**russel**	**cosine**	**ejaccard**	**edice**	**correlation**
research budget					1st quartile = reference category					
2nd quartile	.042 (.055)	.106 (.068)	.017 (.011)	.017 (.011)	-.002 (.005)	.004 (.008)	.009 (.010)	.010 (.010)	.010 (.010)	.006 (.011)
3rd quartile	-.003 (.057)	.048 (.071)	.002 (.011)	.002 (.011)	-.008 (.006)	.021** (.008)	.024** (.011)	.025** (.011)	.025** (.011)	.025** (.012)
4th quartile	.137* (.072)	.346*** (.089)	.051*** (.015)	.051*** (.015)	-.023*** (.006)	.047*** (.010)	.060*** (.013)	.061*** (.014)	.061*** (.014)	.059*** (.015)
academic rank					junior = reference category					
postdoctoral position	.279** (.124)	.827*** (.141)	.068*** (.022)	.069*** (.022)	-.045*** (.008)	.078*** (.014)	.093*** (.018)	.095*** (.018)	.095*** (.018)	.077*** (.018)
assistant professor	.317** (.129)	.972*** (.148)	.079*** (.023)	.079*** (.023)	-.052*** (.009)	.093*** (.015)	.108*** (.019)	.109*** (.019)	.110*** (.019)	.092*** (.020)
full professor	.411*** (.134)	1.175*** (.161)	.079*** (.025)	.080*** (.025)	-.072*** (.011)	.122*** (.017)	.137*** (.021)	.138*** (.021)	.139*** (.021)	.114*** (.022)
institution rank					not ranked = reference category					
tier 3	-.010 (.053)	.072 (.067)	.010 (.010)	.010 (.010)	-.004 (.005)	.005 (.007)	.005 (.010)	.005 (.010)	.005 (.010)	.000 (.011)
tier 2	.094** (.046)	.195*** (.057)	.029*** (.009)	.029*** (.009)	-.012*** (.005)	.023*** (.007)	.032*** (.009)	.032*** (.009)	.032*** (.009)	.029*** (.010)
tier 1	.205*** (.048)	.294*** (.060)	.032*** (.010)	.032*** (.010)	-.011** (.005)	.022*** (.007)	.038*** (.009)	.038*** (.009)	.038*** (.010)	.040*** (.010)
publications	.007*** (.001)	.035*** (.003)	-.001*** (.000)	-.001*** (.000)	-.005*** (.000)	.006*** (.000)	.004*** (.000)	.003*** (.000)	.003*** (.000)	.002*** (.000)
co-authors per article	.059*** (.009)	.090*** (.012)	.003** (.001)	.003** (.001)	-.008*** (.001)	.011*** (.002)	.010*** (.002)	.010*** (.002)	.010*** (.002)	.007*** (.002)
age	.010 (.014)	.037** (.017)	.003 (.003)	.003 (.003)	-.005*** (.001)	.007*** (.002)	.006** (.002)	.006** (.002)	.006** (.002)	.005* (.003)
age^2	-.000 (.000)	-.000** (.000)	-.000 (.000)	-.000 (.000)	.000*** (.000)	-.000*** (.000)	-.000*** (.000)	-.000*** (.000)	-.000*** (.000)	-.000** (.000)
female	.052 (.052)	-.055 (.061)	-.020* (.010)	-.020* (.010)	.003 (.005)	-.011* (.007)	-.015* (.009)	-.016* (.009)	-0.000144	-.015 (.010)
country					Japan = reference category					
Germany	.536*** (.062)	.705*** (.080)	.119*** (.012)	.119*** (.012)	-.062*** (.006)	.122*** (.009)	.162*** (.012)	.163*** (.012)	.163*** (.012)	.170*** (.013)
United Kingdom	.634*** (.062)	.749*** (.077)	.151*** (.012)	.151*** (.012)	-.053*** (.006)	.119*** (.009)	.179*** (.011)	.180*** (.012)	.180*** (.012)	.202*** (.012)
field					engineering = reference category					
biology	.393*** (.054)	.406*** (.067)	-.088*** (.010)	-.088*** (.010)	-.008 (.005)	.003 (.007)	-.049*** (.009)	-.060*** (.009)	-.061*** (.009)	-.074*** (.010)
chemistry	.429*** (.050)	.587*** (.065)	-.105*** (.010)	-.107*** (.010)	.074*** (.005)	-.054*** (.007)	-.117*** (.009)	-.117*** (.009)	-.119*** (.009)	-.068*** (.010)
economics	-.270*** (.066)	-.522*** (.084)	-.087*** (.013)	-.087*** (.013)	-.016** (.007)	-.020** (.010)	-.067*** (.012)	-.080*** (.012)	-.081*** (.012)	-.108*** (.013)
_cons	.132 (.335)	.454 (.398)	.483*** (.065)	.486*** (.065)	1.136*** (.027)	-.175*** (.042)	.048 (.057)	.059 (.058)	.063 (.058)	.141** (.062)
observations	1884	1884	1884	1884	1884	1884	1884	1884	1884	1884
R^2	.31	.63	.22	.22	.76	.67	.42	.41	0.41	0.3

Notes: *** (**,*) indicate a significance level of 1% (5%, 10%).

4.7 Conclusions

Scientific excellence plays a crucial role in research funding and the identification of individual scientific excellence remains a challenge in practice and theory. This is mostly because quality and excellence of scientists are never fully observed but can rather only be approximated, for example by peer review and bibliometric analyses, or content-based analyses. While previous research has provided vast investigations on the identification from peer-review and bibliometric analyses, the potential of content-bases analyses for science evaluations has hardly been addressed and remains far from understood.

The prime goal of the efforts described in this paper was to explore the technical feasibility and plausibility of content-based indicators for identifying scientific excellence of individual scientists. The research question which we address is here was whether text-based similarity between publications of individual scientists and documents of validated knowledge frontiers can be used to evaluate scientific excellence. To answer this question, we conducted a comparative analysis of standard and new indicators of scientific excellence to see if they provide a coherent picture or if they point to perpendicular directions.

The results confirm that document-document similarity between individual scientists' publications and knowledge frontier documents indeed captures scientific excellence. We find that four common research quality indicators (i.e. citations, research budget, academic rank and institution rank) show a positive correlation with the derived text similarity indicators. We interpret these findings as some initial evidence for the idea that content-based analyses based on knowledge frontiers can be valuable for science evaluations when citation measures may be less meaningful. This is potentially the case for younger scholars since their citation numbers had less time to accumulate.

Our analysis provides three new insights to the academic discussion on augmented research evaluations. First, we propose a new empirical method that approximates the scientific proximity between individual scientists and validated knowledge frontiers of excellence. This method rests upon the rarely used contents of scientific communication (publications), rather than on citation counts and their weaknesses (MacRoberts and MacRoberts (1996). Second, our analysis introduces the large scale use of academic prizes and prestigious funding awards for science evaluation. To the best of our knowledge, no previous study has used such knowledge frontier definitions as a benchmark for scientific excellence and science evaluations. Third, we illuminate the feasibility of eight similarity measures for their use in science evaluations. While scholars in information sciences and related fields have begun to characterize their usefulness for information retrieval or document recommender systems, none of them has provided a characterization of such measures for scientific documents. It turns out that some of the utilized similarity measures actually provide mutual information (Jaccard, Dice for the binary case and eJaccard, eDice and

Cosine for the metrics case), while two measure somewhat deviate from this group (i.e. simple matching coefficient and Russel index). We propose that these deviating similarity measure might not be the best choice for deriving the "true" similarity of text documents.

Our study shows that content-based indicators are a valuable source of information which can complement peer review and standard bibliometric indicators. The results suggest that policy makers and administrators may consider such indicators to for research funding allocation and science evaluations.

The precision and validity of this method depends on several critical factors and assumptions. These include a thorough definition of "scientific excellence" and sophisticated text pre-processing. We define scientific excellence by two externally validated knowledge frontiers, in contrast to previous studies which define scientific excellence by imposing a fixed threshold on scientists' research quality distribution (Bonaccorsi et al. 2017; Bornmann 2014). The results suggest that this external definition of excellence provides a robust benchmark, which is independent of arbitrary choices in defining scientific excellence as the top X percentile. Further, we perform a comparison of two different knowledge frontiers throughout this analysis which allows us to verify, at least to some extent, that the method is also compelling in an independent setting.

Co-word analysis critically relies on the pre-processing of text inputs, which consist of discriminating terms rather than trivial terms (e.g. method, result, findings). To ensure that the similarity scores are based on discriminating terms, we tested and analysed different parameter settings suggested by theory, especially Salton and Buckley (1988). The presented findings in this analysis use a configuration of parameters (term-weighting scheme, token size, stop-word lists, and term frequency boundaries), that are best suitable for co-word analysis of scientific documents.

Limitations and future research

Content-based analysis is on the rise in many scientific disciplines. However their use in science evaluation remains scarce. While a large body of research literature on citations exist, a compelling understanding of the value of text in science evaluations is needed. We suggest that more studies on the identification of excellence try to exploit information based on scientific text documents.

Here, we illustrate potential future research directions that emerged throughout this study. For example, the influence of prestigious research grants on scientific excellence remains anecdotal. We encourage future studies to provide empirical evidence for the idea that scientists' ability to obtain competitive funding from prestigious research bodies can indeed signal scientific excellence. This would justify the use of funding indicators for future science evaluations.

We used ERC project descriptions to define a prestigious funding knowledge frontier. Instead of using the project descriptions, future research could also incorporate the publication records of principal investigators, to create text similarity scores. This would result in an extended funding frontier, which is deep, due to the publication portfolio of the principal investigator and also broad due to the large number of ERC projects. In contrast, future research could also define a “placebo” frontier, for example based on low ranked scientific publications or on unrelated documents. It would be interesting to see how such an extended funding frontier, or a placebo frontier relate to our findings.

The choice of similarity measures for use on co-word analysis also remains unclear. Our efforts to compare several measures of text coherence provides an overview of their relative boundaries. However a clear answer to which similarity measure is the best approximation for the “true” similarity of documents, is still needed. Further research might consider simulations, to determine which similarity measure is best suited for co-word analysis of scientific documents.

One important drawback of the proposed method is its insufficiency to deal with lexical ambiguity and variability, for example synonymy, antonymy, homonymy, polysemy, acronyms, negations, alternations, abbreviations, etc. (Hotho et al. 2005). Future research might consider other methods that account for lexical ambiguity (e.g. topic models, part-of speech tagging), to overcome such ambiguity.

The main assumption in this text similarity approach is that sample authors are closer to the knowledge frontier if they share common vocabularies and concepts in the writing of their publications with frontier documents. However several issues could prevent this. Cutting-edge scientists for example might introduce new terms and use a vocabulary that is distant to text documents of current knowledge frontier definitions. Their scientific excellence would remain unrecognized by our method. However, even if new scientific term appear in the literature, they are always embedded in some context. We believe that the text-similarity methods would still provide a meaningful indicator.

References

Aksnes, D. W., and Sivertsen, G. (2004). The effect of highly cited papers on national citation indicators. *Scientometrics*, 59(2), 213-224.

Ayral, H., and Yavuz, S. (2011). An automated domain specific stop word generation method for natural language text classification. In *Innovations in Intelligent Systems and Applications* (INISTA), 2011 International Symposium on, 500-503, IEEE.

Beel, J., Gipp, B., Langer, S., and Breitinger, C. (2016). Research-paper recommender systems: a literature survey. *International Journal on Digital Libraries*, 17(4), 305-338.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.

Börner, K., Chen, C., and Boyack, K. (2003). Visualizing Knowledge Domains. In Blaise Cronin (Ed.), *Annual Review of Information Science & Technology*, Volume 37, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, chapter 5, 179-255.

Bornmann, L. (2014). How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Research Evaluation*, 23(2), 166–173.

Börner, K. (2010). Atlas of science: Visualizing what we know. Cambridge, MA: MIT Press.

Bollacker, K. D., Lawrence, S., and Giles, C. L. (1998, May). CiteSeer: An autonomous web agent for automatic retrieval and identification of interesting publications. In *Proceedings of the second international conference on Autonomous agents* (pp. 116-123). ACM.

Bonaccorsi, A., Cicero, T., Haddawy, P., and Hassan, S. U. (2017). Explaining the transatlantic gap in research excellence. Scientometrics, 110(1), 217-241.

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., and Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, 62(10), 2765-2783.

Boyack, K. W., and Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately?. J*ournal of the Association for Information Science and Technology*, 61(12), 2389-2404.

Catalini, C., Lacetera, N., and Oettl, A. (2015). The incidence and role of negative citations in science. *Proceedings of the National Academy of Sciences*, 112(45), 13823-13826.

Charlton, B. G. (2007a). Measuring revolutionary biomedical science 1992–2006 using Nobel prizes, Lasker (clinical medicine) awards and gairdner awards (NLG metric). *Medical Hypotheses*, 69(1), 1–5.

Charlton, B. G. (2007b). Which are the best nations and institutions for revolutionary science 1987–2006? Analysis using a combined metric of Nobel prizes, fields medals, Lasker awards and turing awards (NFLT metric). Medical Hypotheses, 68(6), 1191–1194. doi:10.1016/j.mehy.2006.12.007.

Chubin, D. E., and Hackett, E. J. (1990). Peerless science: Peer review and US science policy. Suny Press.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science*, 41(6), 391.

Deza, M. M., and Deza, E. (2009). Encyclopedia of distances. In *Encyclopedia of Distances*, 1-583, Springer Berlin Heidelberg.

Ding, Y., Chowdhury, G. G., and Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information processing & management*, 37(6), 817-842.

Ding, Y., Zhang, G., Chambers, T., Song, M., Wang, X., and Zhai, C. (2014). Content-based citation analysis: The next generation of citation analysis. *Journal of the Association for Information Science and Technology*, 65(9), 1820-1833.

Donovan B. (1998). The truth about peer review. *Learned Publishing,* 11(3), 179-184.

Donovan, C., and Butler, L. (2007). Testing novel quantitative indicators of research 'quality', esteem and 'user engagement': An economics pilot study. *Research Evaluation*, 16(4), 231-242.

Doszkocs, T. E., Rapp, B. A., and Schoolman, H. M. (1980). Automated information retrieval in science and technology. *Science*, 208(4439), 25-30.

Frakes, W. B., and Baeza-Yates, R. (1992). Information retrieval: Data structures & algorithms (Vol. 331). Englewood Cliffs, New Jersey: prentice Hall.

Frey, B. S. (2007). Awards as compensation. *European Management Review,* 4(1), 6-14.

Frey, B. S., and Neckermann, S. (2008). Academics appreciate awards-a new aspect of incentives in research.

Frey, B. S., and Neckermann, S. (2010). Awards as signals. CESifo Working Paper Series No. 3229

Jaccard, P. (1912). The distribution of the flora in the alpine zone. *New Phytologist*, 11(2), 37-50.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?. The Review of Economic Studies, 76(1), 283-317.

Garfield, E. (1986). Do Nobel-Prize winners write citation-classics? *Current Contents,* (23), 3-8.

Garfield, E., and Welljams-Dorof, A. (1992). Citation data: their use as quantitative indicators for science and technology evaluation and policy-making. Science and Public Policy, 19(5), 321-327.

Griffiths, T. L., and Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(1), 5228–5235.

Gollapalli, S. D., Mitra, P., and Giles, C. L. (2012). Similar researcher search in academic environments. In *Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries,* pp. 167-170.

Hartigan, J. A. (1967). Representation of similarity matrices by trees. *Journal of the American Statistical Association*, 62(320), 1140-1158.

Hansen, W. L., and Weisbrod, B. (1972). Toward a general theory of awards, or, do economists need a hall of fame? *Journal of Political Economy*, 80(2), 422-431.

He, Q. (1999). Knowledge discovery through co-word analysis. *Library trends,* 48(1), 133-59.

Hicks, D., Breitzman Sr, A., Hamilton, K., and Narin, F. (2000). Research excellence and patented innovation. *Science and Public Policy*, 27(5), 310-320.

Hicks, D., Tomizawa, H., Saitoh, Y., and Kobayashi, S. (2004). Bibliometric techniques in the evaluation of federally funded research in the United States. Research Evaluation, 13(2), 76-86.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences of the United States of America*, 102(46), 16569.

Hotho, A., Nürnberger, A., and Paaß, G. (2005). A brief survey of text mining. In LDV Forum - GLDV Journal for Computational Linguistics and Language Technology, 20(1), 19-62.

Hottenrott, H., and Lawson, C. (2017). Fishing for complementarities: Research grants and research productivity. International Journal of Industrial Organization, 51, 1-38.

Iaria, A., Schwarz, C., and Waldinger, F. (2018). Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science. The Quarterly Journal of Economics, forthcoming.

Inhaber, H., and Przednowek, K. (1976). Quality of research and the Nobel prizes, *Social Studies of Science*, 6, 33–50.

Kostoff, R. N. (1997). The Handbook of Research Impact Assessment. Edition 7. Summer 1997. Office of naval research Arlington VA.

Krovetz, R. (1993). Viewing morphology as an inference process. In *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR'93, 191-202.

Lamont, M. (2009). How professors think. Harvard University Press.

Lam, A. (2011). What motivates academic scientists to engage in research commercialization: ‘Gold’,‘ribbon’or ‘puzzle’? *Research policy*, 40(10), 1354-1368.

Law, J., Bauin, S., Courtial, J., and Whittaker, J. (1988). Policy and the mapping of scientific change: A co-word analysis of research into environmental acidification. *Scientometrics*, 14(3-4), 251-264.

Lee, C. J., Sugimoto, C. R., Zhang, G., and Cronin, B. (2013). Bias in peer review. *Journal of the Association for Information Science and Technology*, 64(1), 2-17.

Lenz, H. J. (2008). Proximities in statistics: Similarity and distance. *Preferences and Similarities*, 161-177.

Leydesdorff, L., and Zaal, R. (1988). Co-words and citations: Relations between document sets and environments. *Informetrics*, 87/88. 105-119.

Lovins, J. B. (1968). Development of a stemming algorithm. *Mech. Translatation and Computational Linguistics*, 11(1-2), 22-31.

MacRoberts, M. H., and MacRoberts, B. R. (1996). Problems of citation analysis. *Scientometrics*, 36(3), 435-444.

Marx, W., and Bornmann, L. (2015). On the causes of subject-specific citation rates in Web of Science. *Scientometrics*, 102(2), 1823-1827.

Manning, C. D., & Schütze, H. (1999). Foundations of statistical natural language processing. MIT press.

Martin, B. R. (1996). The use of multiple indicators in the assessment of basic research, *Scientometrics*, 36(3), 343-362.

Meho, L. I. (2007). The rise and rise of citation analysis. *Physics World*, 20(1), 32.

Merton, R. K. (1968). The Matthew effect in science. Science, 159(3810), 56-63.

Meyer, D., Hornik, K., and Feinerer, I. (2008). Text mining infrastructure in R. *Journal of statistical software*, 25(5), 1-54.

Narin, F. (1976). Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity, 206-219, Washington, DC: Computer Horizons.

Neckermann, S., and Frey, B. S. (2007). Awards as incentives. Institute for Empirical Research in Economics, University of Zurich.

Ide, N. and Veronis, J. (1998). Word Sense Disambiguation: The State of the Art, *Computational Linguistics*, 24(1), 1-40.

Persson, O. (1994). The intellectual base and research fronts of JASSIS 1986–1990. JASSIS, 45(1), 31–38.

Persson, O., Glänzel, W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3), 421-432.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130-137.

Rowland, F. (2002). The peer-review process. *Learned publishing*, 15(4), 247-258.

Salton, G., and Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information processing & management*, 24(5), 513-523.

Saviotti, P. P. (1998). On the dynamics of appropriability, of tacit and of codified knowledge. *Research policy*, 26(7), 843-856.

Schiebel, E. (2012). Visualization of research fronts and knowledge bases by three-dimensional areal densities of bibliographically coupled publications and co-citations. Scientometrics, 91(2), 557-566.

Schmid, H. (1995). Treetagger - a language independent part-of-speech tagger. Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, 43, 28.

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628.

Small, H. (1999). Visualizing science by citation mapping. *Journal of the Association for Information Science and Technology*, 50(9), 799.

Stephan, P. E. (2012). How economics shapes science (Vol. 1). Cambridge, MA: Harvard University Press.

Sternitzke, C., and Bergmann, I. (2009). Similarity measures for document mapping: A comparative study on the level of an individual scientist. *Scientometrics*, 78(1), 113-130.

Tijssen, R., Visser, M., and van Leeuwen, T. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics*, 54(3), 381-397.

Tijssen, R. J. (2003). Scoreboards of research excellence. *Research Evaluation*, 12(2), 91-103.

Tijssen, R., and Kraemer-Mbula, E. (2017). Research excellence in Africa: Policies, perceptions, and performance. *Science and Public Policy*.

Van Leeuwen, T., Visser, M., Moed, H., Nederhof, T., and Van Raan, A. (2003). The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics*, 57(2), 257-280.

Vinkler, P. (2007). Eminence of scientists in the light of the h-index and other scientometric indicators. *Journal of information science*, 33(4), 481-491.

Weinstock, N. (1971), Citation indexes, In: KENT, A. (Ed.). *Encyclopedia of Library and Information Science*, New York: Marcel Dekker, Vol. 5, 1971, pp. 16–41.

Hansen, W. L., and Weisbrod, B. (1972). Toward a general theory of awards, or, do economists need a hall of fame?. *Journal of Political Economy*, 80(2), 422-431.

Wright, B. D. (1983). The economics of invention incentives: Patents, prizes, and research contracts. *The American Economic Review*, 73(4), 691-707.

Yang, Y. (1995). Noise reduction in a statistical approach to text categorization. In *Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 256–263.

Zheng, J., and Liu, N. (2015). Mapping of important international academic awards. Scientometrics, 104(3), 763-791.

Appendix

Survey description

International Science Affiliations Survey 2016

In June to August 2016, we conducted an online survey of corresponding authors in five scientific disciplines: biology, chemistry, engineering, economics & business, and history. In order to construct the sample, we first selected all journals classified by the 2013 Journal Citation Report (Thomson Reuters) as belonging to the five fields and sorted them by eigenfactor score in each of the disciplines. The eigenfactor score is a rating of journal importance based on the number of incoming, journal-weighted citations that enables us to consider journals across all quality spectra. From each quartile of the eigenfactor distribution we randomly drew five journals, 20 in total for each field. As the number of articles in the selected journals was very low for engineering, economics & business and history, we drew additional journals in these fields resulting in 40 journals in engineering, 80 in economics & business, and 40 in history. The process resulted in five samples of journals by field, stratified by eigenfactor score. Historians were not surveyed in the case of Japan due to the low number of articles in WOS listed journals.

Articles published in each of the selected journals and with a reprint address in a university or public research organization in Germany, Japan or the UK were downloaded from the Web of Science (WoS) for the years 2013 to 2015. We retrieved the email addresses of the corresponding authors. After some manual cleaning, we were left with a final list of 8916 corresponding authors.

The survey consisted of four sections and was designed to be completed in 10-20 minutes. The questions sought to discover involvement in multiple affiliations and how and why these affiliations were formed as well as involvement in and the organisation of data sharing. The survey was conducted in German, Japanese and English through the platform LimeSurvey. The emails containing the survey link were sent from the personal email accounts of the principal investigators; they explained how email addresses were collected and provided a survey opt-out function.

Supplementary tables

Table 4.11: Descriptive statistics

Variable	unit	source	median	mean	s.d.	min.	max.
Average similarity scores prize frontier							
jaccard	score	Scopus	0	0.48	0.21	0	1
dice	score	Scopus	1	0.51	0.21	0	1
smc	score	Scopus	1	0.83	0.14	0	1
russel	score	Scopus	0	0.22	0.16	0	1
cosine	score	Scopus	0	0.40	0.18	0	1
ejaccard	score	Scopus	0	0.39	0.18	0	1
edice	score	Scopus	0	0.41	0.18	0	1
correlation	score	Scopus	0	0.43	0.19	0	1
Average similarity scores funding frontier							
jaccard	score	Scopus/Cordis	1	0.62	0.17	0	1
dice	score	Scopus/Cordis	1	0.62	0.17	0	1
smc	score	Scopus/Cordis	1	0.79	0.15	0	1
russel	score	Scopus/Cordis	0	0.31	0.19	0	1
cosine	score	Scopus/Cordis	0	0.47	0.19	0	1
ejaccard	score	Scopus/Cordis	0	0.47	0.19	0	1
edice	score	Scopus/Cordis	0	0.48	0.19	0	1
correlation	score	Scopus/Cordis	0	0.47	0.19	0	1
Publication information							
publications$_{2011\text{-}2016}$	count	Scopus	18.92	22.09	1	237	18.92
citations	count	Scopus	186.70	381.35	1	7332	186.70
citations per publication	fraction	Scopus	7.97	13.67	0	519	7.97
co-authors per publication	fraction	Scopus	5.42	2.98	1	45	5.42
Controls							
age	count	Survey	46.21	10.85	25	88	46.21
female	binary	Survey	0.17	0.37	0	1	0.17
budget	amount	Survey	0.15	4.69	20.83	0	600
junior	binary	Survey	0	0.04	0.20	0	1
postdoc	binary	Survey	0	0.26	0.44	0	1
assistant professor	binary	Survey	0	0.31	0.46	0	1
full professor	binary	Survey	0	0.38	0.49	0	1
institution rank: not ranked	binary	THE Ranking	0	0.36	0.48	0	1
institution rank: Tier 1	binary	THE Ranking	0	0.18	0.38	0	1
institution rank: Tier 2	binary	THE Ranking	0	0.23	0.42	0	1
institution rank: Tier 3	binary	THE Ranking	0	0.24	0.42	0	1
Japan	binary	Survey	0	0.30	0.46	0	1
United Kingdom	binary	Survey	0	0.43	0.50	0	1
Germany	binary	Survey	0	0.27	0.45	0	1
biology	binary	Survey	0	0.27	0.44	0	1
chemistry	binary	Survey	0	0.31	0.46	0	1
economics	binary	Survey	0	0.20	0.40	0	1
engineering	binary	Survey	0	0.22	0.42	0	1

Notes: Number of observations = 1884, Funding variables in million €. THE: Times Higher Education

Table 4.12: Mapping of fields

Sample authors (ISA-Survey 2016)	**Academic prize awardees (Scopus)**	**Prestigious research funding (ERC)**
	Engineering	
Chemical, Thermal and Process Engineering, Computer Science, IT and Electrical and Electronic Engineering, Materials Science and Engineering, Mechanical, Aeronautical and Manufacturing Engineering, Civil and Construction Engineering; Architecture	Materials Science, Engineering, Energy, Computer Science, Chemical Engineering	Information Processing and Information Systems, Information and communication technology applications, Network technologies, Telecommunications, Electronics and Microelectronics, Physical sciences and engineering, Nanotechnology and Nanosciences, Space and satellite research, Aerospace Technology, Materials Technology, Industrial Manufacture, Construction Technology
	Economics/Business	
Arts and the Humanities, Business Administration, Economics	Arts and Humanities, Business, Management and Accounting, Economics, Econometrics and Finance, Decision Sciences, Social Sciences, Psychology	Social sciences and humanities, Business aspects, Economic Aspects, Regional Development
	Biology/Medicine	
Neurosciences, Agriculture, Forestry and Veterinary Medicine, Biological Sciences, Medicine (including Pharmacy, Dentistry and Nursing)	Neuroscience, Agricultural and Biological Sciences, Veterinary, Biochemistry, Genetics and Molecular, Biology, Immunology and Microbiology, Medicine, Pharmacology, Toxicology and Pharmaceutics, Nursing, Dentistry, Health Professions	Agricultural biotechnology, Life Sciences, Biotechnology, Medicine and Health, Medical biotechnology, Healthcare delivery/services
	Chemistry	
Chemistry, Geosciences (including Geography), Mathematics, Physics	Chemistry, Environmental Science, Earth and Planetary Sciences, Mathematics, Physics and Astronomy	Earth Sciences, Environmental Protection, Mathematics and Statistics, Physical sciences and engineering, Materials Technology

Table 4.13: List of Academic Prizes by discipline

Discipline / academic prize	Years considered	Award cycle
Economics and Business[26]		
The Erwin Plein Nemmers prize in economics	2008-2016	biennial
Yrjö Jahnsson Award	2009-2017	biennial
Deutsche Bank Prize in Financial Economics	2007-2015	biennial
BBVA foundation frontiers of knowledge award in economics, finance and management	2012-2016	annual
IZA prize in labor economics	2012-2016	annual
The Stephen A. Ross prize in financial economics	2008-2016	biennial
Bernacer Prize	2012-2016	annual
Leontief Prize	2013-2017	annual
Global economy prize for economics	2013-2017	annual
The Ewing Marion Kauffman prize medal for distinguished research in entrepreneurship	2013-2017	annual
Life Sciences[27]		
Crafoord prize in Biosciences	1999-2015	quadrennial
Darwin Medal	2008-2016	biennial
International Prize for Biology	2012-2016	annual
Louisa-Gross-Horwitz-Preis	2012-2016	annual
Heineken prize for biochemistry and biophysics	2008-2016	biennial
Breakthrough Prize in Life Sciences	2013-2017	annual
TWAS prize in Biology	2012-2016	annual
International cosmos prize	2012-2016	annual
ASBMB–Merck Award	2013-2017	annual
The Danone International Prize for Nutrition	2008-2016	biennial
Chemistry[28]		
Wolf Prize in Chemistry	2013-2017	annual
Priestley Medal	2013-2017	annual
Welch award in chemistry	2012-2016	annual
NAS award in chemical sciences	2013-2017	annual
Faraday lectureship prize	2012-2016	annual
Davy medal	2012-2016	annual
Benjamin Franklin medal in chemistry	2013-2017	annual
Peter Debye award in physical chemistry	2013-2017	annual
Roger Adams award in organic chemistry	2009-2017	biennial
TWAS prize in chemistry	2012-2016	annual
Claude S. Hudson award in carbohydrate chemistry	2009-2017	biennial
Engineering[29]		
Charles Stark Draper Prize	2012-2016	annual
John Fritz Medal	2012-2016	annual
Queen Elisabeth Prize for Engineering	2009-2017	biennial
Kyoto prize in advanced technology	2013-2017	annual
Kavli Prize in Nanoscience	2008-2016	biennial
Faraday Medal	2012-2016	annual
Millennium technology prize	2008-2016	biennial
TWAS prize in engineering sciences	2012-2016	annual
R.H. Wilhelm award in chemical reaction engineering	2012-2016	annual
Alpha Chi Sigma award for chemical engineering	2012-2016	annual
Founders award for outstanding contributions to the field of chemical engineering	2012-2016	annual

[26] Economics, Finance, Macoreconomics
[27] Biology: Bioscience, Biology, Biochemistry, Nutrition
[28] Chemistry: Chemistry, Physical Chemistry, Organic Chemistry
[29] Engineering: Nanoscience, Chemical engineering, Civil engineering, Electrical and Information Engineering, Environmental science and engineering, Materials science and engineering, Mechanical engineering

Andreas Acrivos Award for Professional Progress in Chemical Engineering	2012-2016	annual
Jacques Villermaux medal	1999-2015	quadrennial
Dieter Behrens medal	1997-2013	quadrennial
Freyssinet medal	2002-2014	quadrennial
International award of merit in structural engineering	2013-2017	annual
IABSE prize	2013-2017	annual
Theodore von Karman medal	2013-2017	annual
Fib medal of merit	2012-2016	annual
A.M. Turing Award	2012-2016	annual
IEEE medal of honor	2013-2017	annual
Benjamin Franklin medal in electrical engineering	2013-2017	annual
IEEE edison medal	2013-2017	annual
The Okawa prize	2012-2016	annual
The Knuth prize	2013-2017	annual
Royal Society Milner award	2013-2017	annual
Benjamin Franklin medal in computer and cognitive science	2013-2017	annual
W. Wallace McDowell award	2013-2017	annual
BBVA foundation frontiers of knowledge award in ICT	2012-2016	annual
World technology award in communications technology (for individuals)	2012-2016	annual
World technology award in it software (for individuals)	2012-2016	annual
World technology award in IT hardware (for individuals)	2012-2016	annual
Eni award	2012-2016	annual
The Enrico Fermi award	2010-2014	annual
The global energy prize	2012-2016	annual
World technology award in energy (for individuals)	2012-2016	annual
Tyler prize for environmental achievement	2013-2017	annual
Volvo environment prize	2012-2016	annual
Stockholm water prize	2012-2016	annual
BBVA foundation frontiers of knowledge award in ecology and conservation biology	2012-2016	annual
BBVA foundation frontiers of knowledge award in climate change	2012-2016	annual
Heineken prize for environmental sciences	2008-2016	biennial
The Zayed international prize for the environment	2008-2016	biennial
World technology award in environment (for individuals)	2012-2016	annual
Von Hippel award	2012-2016	annual
MRS medal award	2012-2016	annual
David Turnbull lectureship	2012-2016	annual
Materials Research Society: Outstanding Young Investigator Award	2012-2016	annual
World technology award in materials (for individuals)	2012-2016	annual
Royal society Armourers & Brasiers company prize	2008-2016	biennial
ASME medal	2013-2017	annual
Timoshenko medal	2013-2017	annual
Benjamin Franklin medal in mechanical engineering	2013-2017	annual
Gibbs brothers medal	2003-2017	triennial
Medicine[30]		
Albert Lasker Award for Basic Medical Research	2012-2016	annual
Lasker-DeBakey Clinical Medical Research Award	2012-2016	annual
Canada Gairdner international award	2013-2017	annual
Shaw Prize in Life Science and Medicine	2012-2016	annual
Canada Gairdner global health award	2013-2017	annual
Wolf Prize in Medicine	2013-2017	annual
Kavli Prize in Neuroscience	2008-2016	biennial
The Louis-Jeantet prize for medicine	2013-2017	annual
Robert Koch Preis	2013-2017	annual
Robert Koch Goldmedallie	2013-2017	annual
Lasker-Koshland special achievement award in medical science	2008-2016	biennial
King Faisal international prize for medicine	2013-2017	annual
Paul Ehrlich and Ludwig Darmstaedter prize	2013-2017	annual

[30] Medicine: Biomedicine, Neuroscience, Polyarthritis, Clinical investigation, Cell Biology

Heineken prize for medicine	2008-2016	biennial
Lewis S. Rosenstiel Award	2012-2016	annual
Wiley prize in biomedical sciences	2013-2017	annual
Massry Prize	2012-2016	annual
Pearl Meister Greengard prize	2012-2016	annual
TWAS prize in Biology	2012-2016	annual
Crafoord prize in polyarthrits	2000-2017	quadrennial
J. Allyn Taylor international prize in medicine	2012-2016	annual
Jessie Stevenson Kovalenko Medal	2008-2016	biennial
Judson Daland prize for outstanding achievement in clinical investigation	2008-2014	varying
Tobias Prize	2008-2016	biennial
Albert Lasker Award for Basic Medical Research	2012-2016	annual

Table 4.14: Correlations with similarity scores based on the prize frontier

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	# articles	1.00																		
2	# citations	0.68*	1.00																	
3	citation per article	0.12*	0.38*	1.00																
4	co-authors	0.23*	0.41*	0.17*	1.00															
5	age	0.24*	0.22*	-0.01	0.11*	1.00														
6	postdoctoral position	-0.11*	-0.20*	-0.03	-0.03	-0.25*	1.00													
7	assistant professor	-0.13*	-0.06*	0.04	0.09*	-0.30*	-0.13*	1.00												
8	full professor	-0.13*	-0.10*	-0.04	-0.08*	-0.24*	-0.14*	-0.40*	1.00											
9	budget	0.29*	0.24*	0.01	0.00	0.61*	-0.17*	-0.47*	-0.53*	1.00										
10	tier 3	-0.09*	-0.12*	-0.03	0.01	0.04	0.01	0.08*	-0.09*	0.00	1.00									
11	tier 2	0.03	0.00	-0.03	-0.04	0.01	0.01	-0.04	0.00	0.03	-0.35*	1.00								
12	tier 1	0.02	0.03	-0.01	-0.03	-0.03	-0.00	-0.07*	0.05	0.01	-0.41*	-0.25*	1.00							
13	jaccard$_{pri}$	0.05	0.11*	0.07*	0.05	-0.03	-0.02	0.00	0.05	-0.04	-0.41*	-0.26*	-0.30*	1.00						
14	dice$_{pri}$	-0.15*	-0.17*	-0.03	-0.09*	-0.18*	0.22*	0.17*	0.04	-0.28*	-0.00	-0.01	-0.00	0.02	1.00					
15	smc$_{pri}$	-0.10*	-0.03	0.06*	-0.10*	-0.11*	0.02	0.10*	-0.05	-0.05	-0.06	-0.01	0.07*	-0.00	-0.33*	1.00				
16	russel$_{pri}$	0.11*	0.07*	0.00	0.07*	0.07*	-0.11*	-0.05	-0.00	0.10*	0.05	0.01	-0.06*	-0.00	-0.34*	-0.34*	1.00			
17	cosine$_{pri}$	0.14*	0.13*	-0.04	0.12*	0.21*	-0.12*	-0.22*	0.01	0.23*	0.01	0.00	-0.00	-0.01	-0.33*	-0.33*	-0.34*	1.00		
18	ejaccard$_{pri}$	0.08*	0.14*	0.05	0.02	-0.05	-0.02	-0.05	0.09*	-0.03	-0.84*	-0.13*	0.30*	0.77*	0.01	0.05	-0.04	-0.01	1.00	
19	edice$_{pri}$	0.62*	0.71*	0.13*	0.22*	0.23*	-0.21*	-0.14*	-0.11*	0.31*	-0.14*	0.00	0.06*	0.10*	-0.15*	0.07*	0.05	0.03	0.15*	1.00
20	correlation$_{pri}$	0.62*	0.71*	0.13*	0.22*	0.22*	-0.21*	-0.14*	-0.10*	0.31*	-0.14*	0.00	0.06*	0.10*	-0.15*	0.07*	0.04	0.03	0.15*	1.00*

Note: 1884 observations.

Table 4.15: Correlations with similarity scores based on the funding frontier

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	# articles	1.00																		
2	# citations	0.68*	1.00																	
3	citation per article	0.12*	0.38*	1.00																
4	co-authors	0.23*	0.41*	0.17*	1.00															
5	age	0.24*	0.22*	-0.01	0.11*	1.00														
6	postdoctoral position	-0.11*	-0.20*	-0.03	-0.03	-0.25*	1.00													
7	assistant professor	-0.13*	-0.06*	0.04	0.09*	-0.30*	-0.13*	1.00												
8	full professor	-0.13*	-0.10*	-0.04	-0.08*	-0.24*	-0.14*	-0.40*	1.00											
9	budget	0.29*	0.24*	0.01	0.00	0.61*	-0.17*	-0.47*	-0.53*	1.00										
10	tier 3	-0.09*	-0.12*	-0.03	0.01	0.04	0.01	0.08*	-0.09*	0.00	1.00									
11	tier 2	0.03	0.00	-0.03	-0.04	0.01	0.01	-0.04	0.00	0.03	-0.35*	1.00								
12	tier 1	0.02	0.03	-0.01	-0.03	-0.03	-0.00	-0.07*	0.05	0.01	-0.41*	-0.25*	1.00							
13	jaccard$_{erc}$	0.05	0.11*	0.07*	0.05	-0.03	-0.02	0.00	0.05	-0.04	-0.41*	-0.26*	-0.30*	1.00						
14	dice$_{erc}$	-0.15*	-0.17*	-0.03	-0.09*	-0.18*	0.22*	0.17*	0.04	-0.28*	-0.00	-0.01	-0.00	0.02	1.00					
15	smc$_{erc}$	-0.10*	-0.03	0.06*	-0.10*	-0.11*	0.02	0.10*	-0.05	-0.05	-0.06	-0.01	0.07*	-0.00	-0.33*	1.00				
16	russel$_{erc}$	0.11*	0.07*	0.00	0.07*	0.07*	-0.11*	-0.05	-0.00	0.10*	0.05	0.01	-0.06*	-0.00	-0.34*	-0.34*	1.00			
17	cosine$_{erc}$	0.14*	0.13*	-0.04	0.12*	0.21*	-0.12*	-0.22*	0.01	0.23*	0.01	0.00	-0.00	-0.01	-0.33*	-0.33*	-0.34*	1.00		
18	ejaccard$_{erc}$	0.08*	0.14*	0.05	0.02	-0.05	-0.02	-0.05	0.09*	-0.03	-0.84*	-0.13*	0.30*	0.77*	0.01	0.05	-0.04	-0.01	1.00	
19	edice$_{erc}$	-0.18*	0.13*	0.08*	-0.04	-0.15*	0.01	0.08*	0.04	-0.11*	-0.10*	-0.03	0.07*	0.07*	0.08*	0.20*	-0.12*	-0.16*	0.12*	1.00
20	correlation$_{erc}$	-0.17*	0.13*	0.08*	-0.04	-0.15*	0.01	0.08*	0.03	-0.10*	-0.10*	-0.03	0.07*	0.08*	0.08*	0.20*	-0.12*	-0.16*	0.12*	1.00*

Note: 1884 observations.

Table 4.16: Correlations with similarities based on the prize frontier

Germany		1	2	3	4	5	6	7	8
1	# articles	1.00							
2	# citations	0.78*	1.00						
3	citation per article	0.06	0.30*	1.00					
4	co-authors	0.24*	0.22*	0.11	1.00				
5	age	0.30*	0.19*	0.01	0.18*	1.00			
6	pos_new	0.34*	0.26*	0.02	-0.07	0.52*	1.00		
7	budget	0.31*	0.18*	0.00	0.14*	0.24*	0.32*	1.00	
8	orgrank	0.08	0.08	-0.00	-0.06	-0.00	0.08	0.01	1.00
9	jaccard	-0.31*	-0.23*	0.06	0.01	-0.18*	-0.10	-0.09	-0.05
10	dice	-0.31*	-0.23*	0.06	0.02	-0.17*	-0.09	-0.09	-0.05
11	smc	-0.83*	-0.63*	-0.06	-0.29*	-0.34*	-0.43*	-0.32*	-0.08
12	russel	0.78*	0.60*	0.08	0.36*	0.31*	0.42*	0.34*	0.05
13	cosine	0.50*	0.40*	0.12*	0.30*	0.17*	0.31*	0.25*	0.02
14	ejaccard	0.49*	0.38*	0.12*	0.30*	0.16*	0.30*	0.25*	0.02
15	edice	0.49*	0.38*	0.12*	0.30*	0.16*	0.30*	0.25*	0.02
16	correlation	0.32*	0.28*	0.14*	0.28*	0.06	0.17*	0.21*	-0.00

Notes: * indicates a significance level of 1%, N = 561.

Japan		1	2	3	4	5	6	7	8
1	# articles	1.00							
2	# citations	0.81*	1.00						
3	citation per article	0.32*	0.56*	1.00					
4	co-authors	0.22*	0.18*	0.30*	1.00				
5	age	0.25*	0.12*	0.03	0.15*	1.00			
6	pos_new	0.26*	0.15*	0.03	-0.00	0.60*	1.00		
7	budget	0.40*	0.45*	0.19*	0.10*	0.16*	0.18*	1.00	
8	orgrank	0.17*	0.16*	0.13*	0.05	-0.06	0.01	0.10*	1.00
9	jaccard	-0.06	-0.08	0.07	0.00	-0.11*	-0.06	0.01	0.18*
10	dice	-0.05	-0.08	0.07	0.01	-0.10*	-0.06	0.01	0.18*
11	smc	-0.82*	-0.56*	-0.27*	-0.35*	-0.27*	-0.27*	-0.33*	-0.20*
12	russel	0.77*	0.54*	0.31*	0.34*	0.25*	0.25*	0.33*	0.24*
13	cosine	0.51*	0.35*	0.26*	0.22*	0.12*	0.15*	0.25*	0.26*
14	ejaccard	0.50*	0.34*	0.26*	0.21*	0.11*	0.14*	0.24*	0.26*
15	edice	0.50*	0.34*	0.26*	0.21*	0.11*	0.14*	0.24*	0.26*
16	correlation	0.36*	0.26*	0.25*	0.14*	0.04	0.07	0.19*	0.24*

Notes: * indicates a significance level of 1%, N=809.

United Kingdom		1	2	3	4	5	6	7	8
1	# articles	1.00							
2	# citations	0.74*	1.00						
3	citation per article	0.18*	0.58*	1.00					
4	co-authors	0.24*	0.39*	0.43*	1.00				
5	age	0.20*	0.14*	0.01	0.00	1.00			
6	pos_new	0.26*	0.22*	0.04	-0.06	0.63*	1.00		
7	budget	0.13*	0.15*	0.13*	0.09	0.09	0.12*	1.00	
8	orgrank	0.03	0.13*	0.19*	0.11	-0.04	-0.06	-0.02	1.00
9	jaccard	-0.16*	-0.10	0.12*	-0.03	-0.05	0.03	-0.03	0.04
10	dice	-0.16*	-0.10	0.12*	-0.02	-0.05	0.03	-0.03	0.04
11	smc	-0.77*	-0.61*	-0.21*	-0.28*	-0.27*	-0.38*	-0.11	-0.03
12	russel	0.71*	0.63*	0.30*	0.33*	0.25*	0.37*	0.13*	0.07
13	cosine	0.48*	0.46*	0.30*	0.26*	0.17*	0.29*	0.09	0.11
14	ejaccard	0.48*	0.46*	0.30*	0.27*	0.16*	0.28*	0.09	0.11
15	edice	0.48*	0.46*	0.30*	0.27*	0.16*	0.28*	0.09	0.11
16	correlation	0.33*	0.37*	0.32*	0.24*	0.10	0.20*	0.08	0.12*

Notes: * indicates a significance level of 1%, N= 514.

Supplementary figures

Figure 4.6: Quantile-Quantile Plots

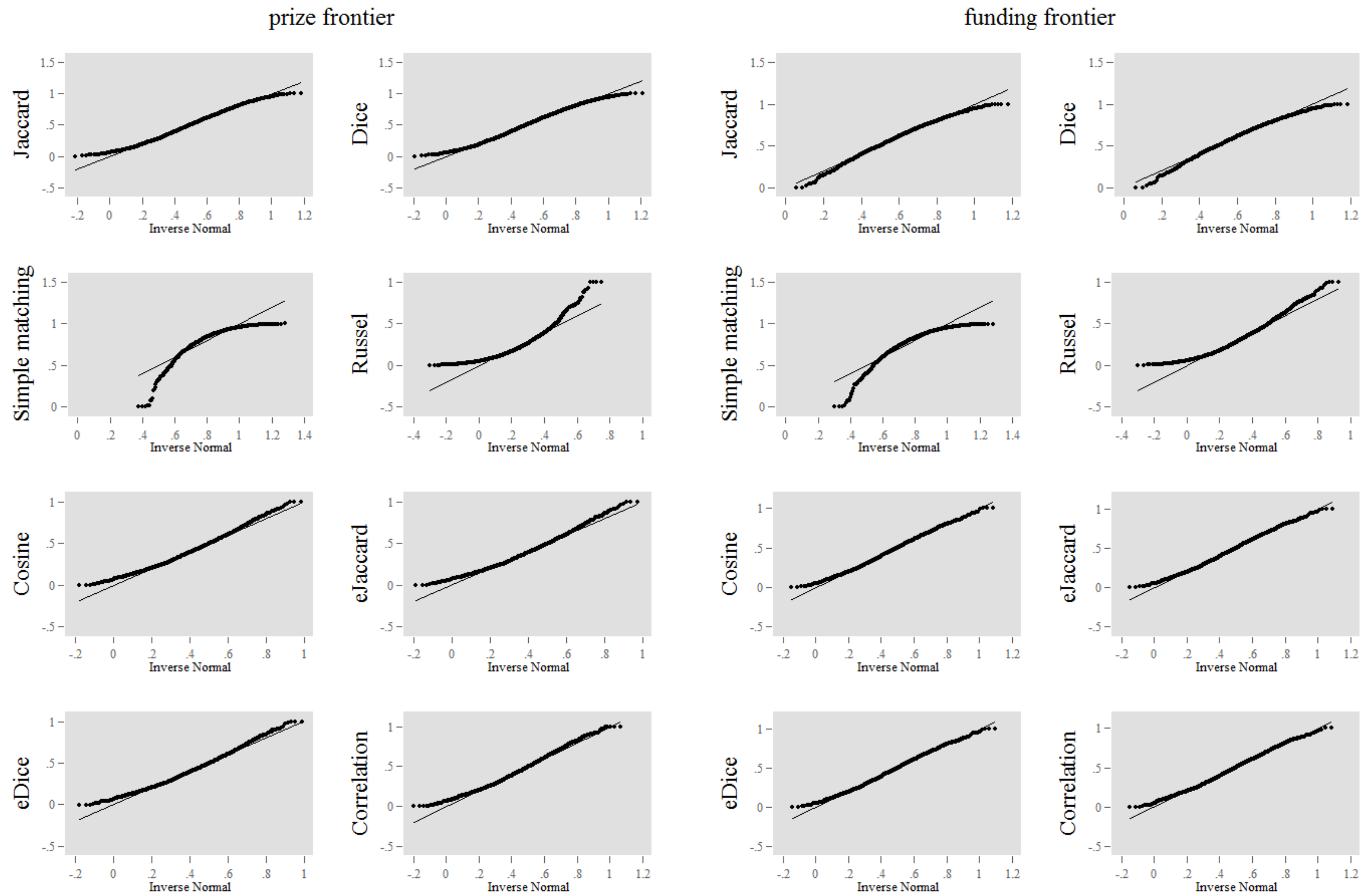


Figure 4.7: Estimated distributions of normalized average similarity scores

(N=1884, unigram, maximum term frequency: 3, 10%)

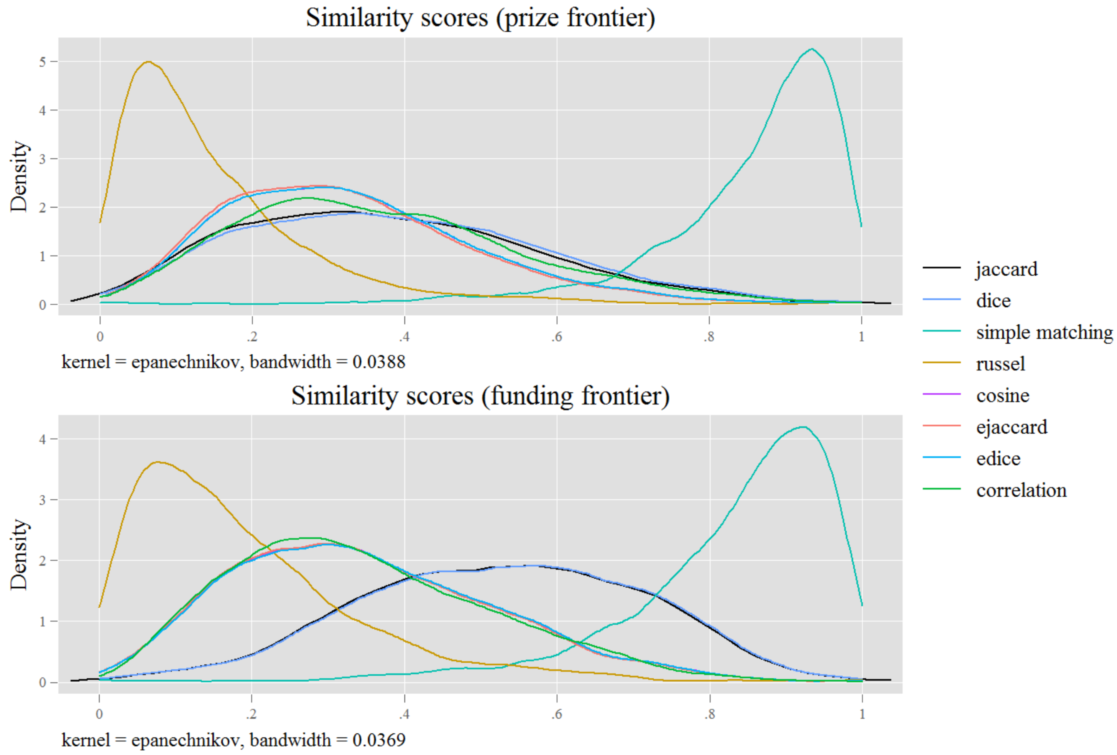


Figure 4.8: Estimated distributions of normalized average similarity scores

(N=1884, bigram, maximum term frequency boundary: 3, 10%)

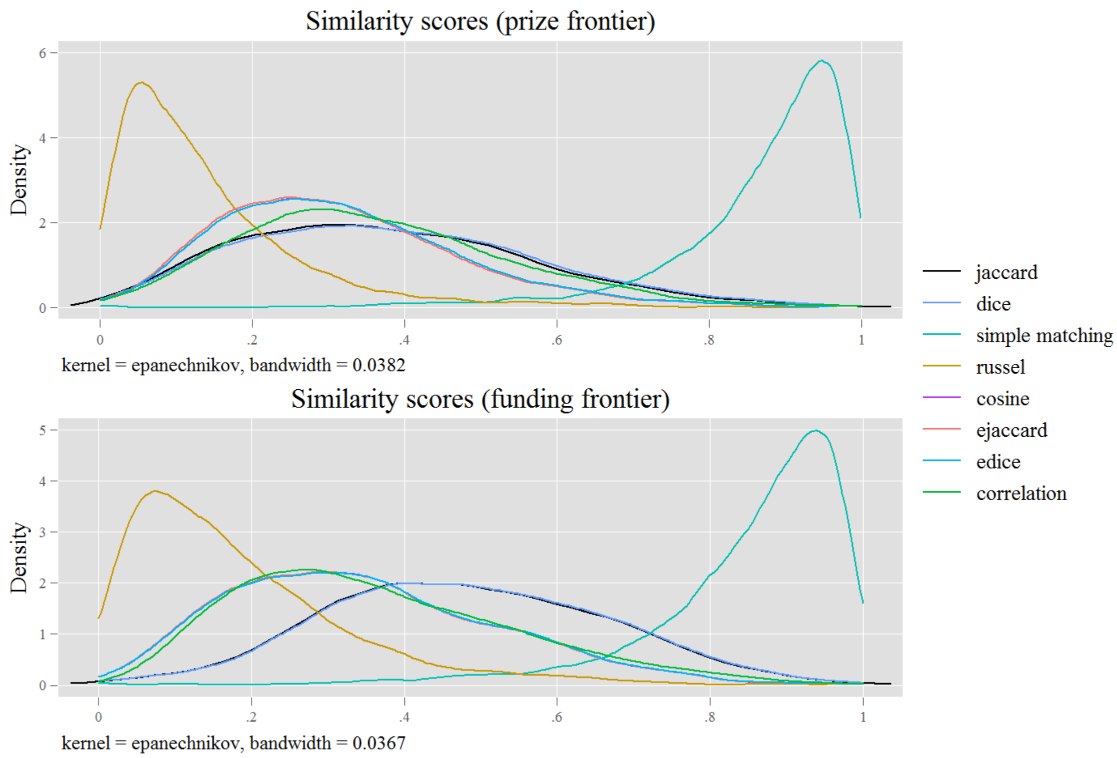


Figure 4.9: Estimated distributions of normalized average similarity scores (N=1884, bigram, maximum term frequency boundary: 3, 33%)

5. Summary and future research

The economic impact of publicly funded research is of great concern to economists and science policy makers. Public research is a key provider of new knowledge, new methods and instruments, and a skilled workforce, which are essential inputs to R&D and innovation in the business sector. To reap the benefits from public research in terms of these economic outcomes, governments invest extensively in their public knowledge infrastructures, which include the networks of people and institutions that generate, maintain, and spread knowledge to boost learning, innovation, and economic growth. Therefore, it is essential to get a deeper understanding of how, how much, and what kind of knowledge is produced in public research institutions, how it flows to extramural users and how this knowledge has a real economic impact.

This thesis provides three essays that characterize such knowledge externalities of publicly funded research. The research presented in this dissertation contributes to the literature primarily by shedding light on two knowledge transfer channels: Academic Entrepreneurship and Academic Consulting. It further provides a new method for the identification of scientific excellence.

Chapter 2 described the relationship between the innovation performance of new technology-based firms (NTBFs) and knowledge interactions with public research institutions in the German context. For a sample of 1708 NTBFs, it is shown that the majority of the sampled firms maintain some form of contacts to PRIs, that these firms interact more often occasionally rather than continuously, and that innovative firms maintain more often and more diverse knowledge interactions with PRIs, compared to non-innovative firms. By the use of probit estimations with Heckman correction, we show that informal and formal interactions can be associated with an increase in NTBFs' innovation probability by 11 and 15%, respectively. The results confirm that R&D-active firms benefit more from knowledge interactions with PRIs. Firms without internal R&D benefit relatively more from informal interaction than those with R&D, particularly when they engage in formal interactions as well. It follows from this discussion that public research is a popular way for firms to acquire complementary knowledge, in order to introduce radically new products and services to the market. To ultimately improve national innovation capacity, the results suggest public policy to further support and facilitate direct knowledge interactions with PRIs, especially on a continuous basis. While these results are interesting, and it is plausible that they generalize to other institutional settings outside of Germany, future work on the evaluation of the impact of knowledge interactions may study these relationships over a longer time horizon to derive stronger conclusions regarding the use and impact of public research as firms mature. In-house basic research and external knowledge sourcing in high and medium-high technology sectors may take up to five years to materialize in an innovation (Higón 2016). Since data on direct knowledge interactions between firms and public research institutions is hardly available, scholars may collect and compare more detailed information of direct knowledge interactions in future research. These include for example details on the time amounts spent on knowledge interactions, their contractual basis, potential monetary compensation, information on the initialization of knowledge interactions, and firms' innovation outcomes.

In chapter 3, we asked whether increased engagement in public or private sector consulting comes at the costs of reduced research performance, or exit from academic work. The analysis investigated the effect of scientists' consulting activity on their subsequent publication and citation numbers in the German context. Estimating probit models that take into account the selection into consulting, the results provide no support for a negative impact of consulting on research disclosure (publication numbers), but a negative effect on average citations per paper in the case of public consulting and particularly by junior and senior researchers. Furthermore, there is some evidence that high amounts of private consulting increases scientists' likelihood to cease publishing (temporarily). The presented analysis suggests that academics that engage in consulting are important knowledge brokers that provide advice to public and private sectors. Yet, this may come at the cost of lower quality research output or the exit from academic publishing. Field differences in the results discussed in chapter 3, however, suggest that consulting does not harm research performance in fields where division of labor is more common, team sizes are larger, and where senior researchers prefer to take the role of science communicators, while more junior researchers conduct the actual research projects. In addition, time distributions of academic scientists have been shown to be useful in the identification of research output effects from commercialization and academic engagement. Future studies might consider the time devoted to alternative knowledge transfer channels for analyzing their detrimental (substitutional) or even positive (complementary) effects on research output. For the case of academic consulting, it might also be interesting to validate previous findings in different countries or institutional settings. Furthermore, a replication of this study for a longer time period would be desirable to account for changes in consulting intensity over time and to differentiate consulting-related effects from general career preferences.

Chapter 4 investigated whether text-based similarity between publications of individual scientists and publications of award or ERC grant winners can be used to evaluate scientific excellence of individual scientists. For a sample of 1884 scientists in three countries (Germany, Japan, and United Kingdom), and in four fields (biology, chemistry, economics and engineering), we calculated average text similarity scores between their scientific documents and documents from two frontier research definitions (scientific prizes and prestigious research funding). The obtained text-similarity indicators had then been compared to other research quality indicators, which are usually associated with closeness to the knowledge frontier, including citation counts, research budgets, academic ranks, and institution ranks. The comparison is carried out by correlation and regression analyses of these potential quality indicators. The findings suggest that document-document similarity between individual scientists' publications and knowledge frontier documents indeed captures to some extent scientific excellence. All four research quality indicators (i.e. citations, research budget, academic rank and institution rank) show a positive correlation with the derived text similarity indicators. One may interpret these findings as some initial evidence for the idea that content-based analyses that rely on knowledge frontiers can be valuable for science evaluations, especially when citation measures may be less meaningful. This is potentially the case for younger scholars since their citation numbers had less time to accumulate. In the process of text analysis of scientific documents it is crucial to retain only words that are discriminative with respect to the scientific contents of a document. While this study is based on co-word analysis on the term level that uses somewhat crude term weighting schemes, alternative methods that overcome lexical ambiguity (i.e. topic models, Blei and Lafferty 2007) or retain only technical terms (Judea et al. 2014) could be useful to identify excellent scientists. Another concern in text-similarity approaches is the choice of appropriate similarity measures. While a plethora of mathematical relationships between term vectors can be calculated (e.g. Jaccard, Cosine, Correlation etc.), it remains unclear which similarity measure is most suitable and precise for scientific documents. Future research could characterize similarity measures with simulated documents that have specific predefined features (e.g. properties of scientific documents) for a better understanding of their significance in content-based analyses.

The aim of this dissertation was to provide a better understanding of how new knowledge originating in public research institutions is produced, disseminated and how it has an economic impact on nations' innovation capacity.

It was shown that public research indeed exhibits positive externalities to the business sector. Academic start-ups but also non-academic start-ups regularly tap new knowledge from public research to complement their own research and development activities. For one knowledge dissemination channel, namely direct knowledge interactions between firms and public research institutions, it was shown that public research improves firms' capacity to introduce radically new products and services to the market. The resulting key suggestion for research policy is to broadly encourage and support direct knowledge interactions with public research, at best on a continuous basis, to maximize the economic impact of public investment in public knowledge infrastructures.

To provide a more complete picture of research externalities, one must also take into account possible negative effects from knowledge transfer activities. Previous research suggested that knowledge transfer activity by academic scientists might impede their individual research productivity (Manjarrés-Henrìquez et al. 2009, Rentocchini et al. 2014). Although the study in chapter 3 finds a slightly reduced research impact (citations) of academic consultants, and an increased probability to exit academic publishing at high levels of consulting, these effects are small compared to the benefits that arise from allowing scientists to provide expertise to public or private sector clients. The exit of junior researchers after receiving their doctorate degree may even be desirable as a way of knowledge transfer between sectors (researcher mobility), and consulting experience may guide researchers to their optimal career paths. As is implemented in many European countries, research policy is well-advised to keep providing soft and flexible regulation, and to monitor the amount of time spent on consulting activities and researchers' engagement in knowledge and technology transfer activities.

While codified knowledge (namely publications) from public research has become widely available in online repositories in recent decades, the new bottleneck which firms face in utilizing public research is access to tacit or implicit knowledge. The first two studies presented in this dissertation emphasize the importance of personal interaction to transmit such implicit knowledge, which cannot be transferred without the active involvement of the knowledge carrier (c.f. the natural excludability of knowledge, Zucker et al. 1998).

It is further shown that identification of excellent scientists is possible, not only through expert ratings (peer review) and bibliometric indicators (publication and citation counts), but also from the publication text itself through content-based analyses combined with knowledge frontier definitions. Due to the high but untapped potential that lies in scientific text documents, it is likely that future evaluations will increase the use of content-based indicators to achieve more efficient research resource allocations, especially when citation measures are less meaningful.

In summary, this dissertation reiterates that public research plays a key role in the knowledge economy, where economic growth is driven by know-how, technology, and innovation. Therefore, governments should retain the high levels of investments in their public knowledge infrastructures. To maximize their economic returns, they should encourage and support the dissemination of research findings through all possible knowledge transfer channels, including direct knowledge interactions and academic consulting, researcher mobility, markets for knowledge, and public or open access. It is emphasized that the active involvement of scientists in knowledge transfer activities is crucial for firms to translate theoretical findings into practical applications. Science governance should therefore encourage and incentivize scientists to regularly leave the ivory tower in order to share scientific knowledge, and to receive new inspiration from the world of technology.

References

Aghion, P., and Howitt, P. W. (2008). The economics of growth. MIT press.

Almus, M. and E. A. Nerlinger (1999). Growth of new technology-based firms: Which factors matter?. *Small Business Economics*, 13(2), 141-154.

Amara, N., Landry, R. and Halilem, N. (2013), Faculty consulting in natural sciences and - engineering: Between Formal and Informal Knowledge Transfer. *Higher Education* 65(3), 359–384.

Antonelli, C. (1999). The evolution of the industrial organisation of the production of knowledge. *Cambridge Journal of Economics*, 23(2), 243–260.

Arora, A., and A. Gambardella (2010), The market for technology. In *Handbook of the Economics of Innovation*, **1**, 641-678.

Arrow, K. (1962), Economic welfare and the allocation of resources for invention. In *The Rate and Direction of Inventive Activity*, edited by R. R. Nelson, Princeton University Press.

Autio, E., and H. Yli-Renko (1998). New, technology-based firms in small open economies - an analysis based on the Finnish experience. *Research Policy*, 26(9), 973-987.

Baum, J. A., Calabrese, T., and B. S. Silverman (2000). Don't go it alone: Alliance network composition and startups' performance in Canadian biotechnology. *Strategic Management Journal*, 21(3), 267-294.

Beise, M., and H. Stahl (1999). Public research and industrial innovations in Germany. *Research Policy*, 28(4), 397-422.

Bekkers, R., and Bodas Freitas, I.M. (2008), Analysing knowledge transfer channels between universities and industry: to what degree do sectors also matter? *Research Policy* 37(10), 1837–1853.

Blei, D. M., and Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, 17-35.

Chubin, D. E., and Hackett, E. J. (1990). Peerless science: Peer review and US science policy. Suny Press.

Cohen, W. M., and D. A. Levinthal (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1), 128-152.

Cohen, W. M., Nelson, R. R., and J. P. Walsh (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science*, 48(1), 1–23.

Dasgupta, P. and P. David (1994). Towards a new economics of science. *Research Policy*, 23(5), 487-521.

Diamond, A. M. (1996). The economics of science. *Knowledge and Policy*, 9(2-3), 6-49.

Donovan, C. (2007). The qualitative future of research evaluation. *Science and Public Policy*, 34(8), 585-597.

Dosi, G. (1988). The Nature of the innovative process, *Technical change and economic theory*, 2, 590-607.

Edwards, P. N. (2010). A vast machine: Computer models, climate data, and the politics of global Warming. MIT Press, Cambridge, MA.

Foray, D., and Lissoni, F. (2010). University research and public–private interaction. In *Handbook of the Economics of Innovation, (*1), 275-314, North-Holland.

Freeman, C., and Soete, L. (1997). The economics of industrial innovation. Psychology Press.

Griliches, Z. (1992). The search for R&D spillovers. *Scandinavian Journal of Economics*, 94, 29–47.

Grimpe, C., and K. Hussinger (2013). Formal and informal knowledge and technology transfer from academia to industry: Complementarity effects and innovation performance. *Industry and Innovation*, 20(8), 683-700.

Godin, B., and Y. Gingras (2000). The place of universities in the system of knowledge production. *Research Policy*, 29(2), 273-278.

Haucap, J. and Thomas, T. (2014). Wissenschaftliche Politikberatung: Erreicht der Rat von Ökonomen Politik und Öffentlichkeit? *Wirtschaftsdienst*, 94(3), 180–186.

Hicks, D., Breitzman Sr, A., Hamilton, K., and Narin, F. (2000). Research excellence and patented innovation. *Science and Public Policy*, 27(5), 310-320.

Higón, D. A. (2016). In-house versus external basic research and first-to-market innovations. *Research Policy*, 45(4), 816-829.

Jaffe, A. B. (1989). Real effects of academic research. *The American Economic Review*, 79(5), 957-970.

Jaffe, A. B. (1998). The importance of “spillovers” in the policy mission of the advanced technology program. *Journal of Technology Transfer*, 23(2), 11-19.

Judea, A., Schütze, H., and Brügmann, S. (2014). Unsupervised training set generation for automatic acquisition of technical terminology in patents. In *Proceedings of COLING 2014, the 25th international conference on computational linguistics: Technical Papers* (pp. 290-300).

Kline, S. and N. Rosenberg (1986). An overview of innovation. In Landau, R. (ed.), *The Positive Sum Strategy. Harnessing Technology for Economic Growth*, 275–306.

Lam, A. (2011). What motivates academic scientists to engage in research commercialization:‘Gold’,‘ribbon’or ‘puzzle’?. *Research policy*, 40(10), 1354-1368.

Lee, Y. S. (1996). ‘Technology transfer’ and the research university: a search for the boundaries of university-industry collaboration. *Research policy*, 25(6), 843-863.

Lee, C. J., Sugimoto, C. R., Zhang, G., and Cronin, B. (2013). Bias in peer review. *Journal of the Association for Information Science and Technology*, 64(1), 2-17.

Manjarrés-Henríquez, L., Gutiérrez-Gracia, A., Carrión-García, A., and Vega-Jurado, J. (2009). The effects of university–industry relationships and academic research on scientific performance: Synergy or substitution? *Research in Higher Education* 50(8), 795-811.

Mansfield, E. (1968), The economics of technological change. New York: W. W. Norton and Co.

Mansfield, E. (1991). Academic research and industrial innovation. *Research Policy*, 20(1), 1-12.

Mansfield, E. (1995). Academic research underlying industrial innovations: Sources, characteristics, and financing. *The Review of Economics and Statistics*, 77(1), 55-65.

Martin, S., and J. T. Scott (2000). The nature of innovation market failure and the design of public support for private innovation. *Research policy*, 29(4), 437-447.

May, R. M. (1997). The scientific wealth of nations. *Science*, 275(5301), 793-796.

Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science*, 159(3810), 56-63.

Mitchell, J.E. and Rebne, D.S. (1995). Nonlinear effects of teaching and consulting on academic research productivity. *Socio-Economic Planning Sciences*, 29(1), 47–57.

Nelson, R. R. (1959). The economics of invention: A survey of the literature. *Journal of Business* 32(2), 101-127.

OECD (2015). “OECD Science, Technology and Industry Scoreboard 2015: Innovation for growth and society”, OECD Publishing, Paris, http://doi.org/http://dx.doi.org/10.1787/sti_scoreboard-2015-en

OECD. (2016). OECD Science, Technology and Innovation Outlook 2016, OECD Publishing, Paris, http://doi.org/10.1787/sti_in_outlook-2016-6-en

OECD (2018), Main Science and Technology Indicators, OECD Publishing, Paris. (retrieved from stats.oecd.org).

Partha, D., & David, P. A. (1994). Toward a new economics of science. *Research policy*, 23(5), 487-521.

Perkmann, M. and Walsh, K. (2008). Engaging the scholar: Three types of academic consulting and their impact on universities and industry, *Research Policy*, 37(10), 1884–1891.

Rebne, D.S. (1989). Faculty consulting and scientific knowledge: A traditional university-Industry linkage. *Educational Administration Quarterly* 25, 338–357.

Rentocchini, F., Manjarrés-Henrìquez, L., D’Este, P. and Grimaldi, R. (2014). The relationship between academic consulting and research performance: Evidence from five Spanish universities. *International Journal of Industrial Organization* 32(1), 70–83.

Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic perspectives*, 8(1), 3-22.

Rowland, F. (2002). The peer-review process. *Learned publishing*, 15(4), 247-258.

Salter, A. J., and B. R. Martin (2001). The economic benefits of publicly funded basic research: a Critical Review. *Research Policy*, 30(3), 509-532.

Stephan, P. E. (1996). The economics of science. *Journal of Economic Literature*, 34(3), 1199-1235.

Stinchcombe, A. L. (1965). Social structure and organizations. In J. G. March (ed.), *Handbook of Organizations*, Rand McNally, Chicago, IL, pp. 142–193.

Storey, D. J., and B. S. Tether (1998). New technology-based firms in the European Union: an introduction. *Research Policy*, 26(9), 933-946.

Tijssen, R. J. (2006). Universities and industrially relevant science: Towards measurement models and indicators of entrepreneurial orientation. *Research Policy*, 35(10), 1569-1585.

Van Leeuwen, T., Visser, M., Moed, H., Nederhof, T., and Van Raan, A. (2003). The Holy Grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics*, 57(2), 257-280.

Zucker, L. G., Darby, M. R., and J. Armstrong (1998). Geographically localized knowledge: Spillovers or markets. *Economic Inquiry* 36(1), 65–86.

List of Figures

List of Tables