J Technol Transf (2022) 47:673–698
DOI 10.1007/s10961-017-9606-8

The evaluation of the Austrian START programme: an impact analysis of a research funding programme using a multi-method approach

Sarah Seus[1] · **Susanne Bührer**[1,2]

Published online: 4 August 2017

Abstract The following article presents and discusses the approach and findings of a recently conducted evaluation study of the Austrian START programme. The START programme is one of Austria's most prestigious research grants for individual researchers at post-doctoral level and provides the grantee with up to 1.2 million Euro for up to 5 years. The programme's aims are twofold: supporting excellent research and qualifying the grantee for a (permanent) senior research position in the research system. The article discusses the effects of the programme and focuses especially on the impacts on the grantees as main beneficiaries. In particular, the scientific output of the grantees and their career development is investigated. Furthermore, the analysis of the indirect beneficiary groups and the analysis of the system in which the START programme is placed, aims at answering the questions whether and how the START programme has contributed to strengthening the capabilities of the Austrian science system. The evaluation uses a control group approach to quantify the effects on the grantees. In order to counterbalance the weaknesses of traditional quantitative impact analysis and to obtain a deeper understanding of the mechanisms of the effects of the funding, the evaluation was complemented by further evidence of a qualitative and quantitative nature.

Keywords Programme evaluation · Impact analysis · Science policy · Research funding · Mixed methods approach · Bibliometric analysis

JEL Classification 03 · 038 · I2 · I23 · I28

✉ Sarah Seus
sarah.seus@isi.fraunhofer.de

[1] Fraunhofer ISI, Breslauer Str. 48, 76139 Karlsruhe, Germany

[2] Eva Heckl KMU Forschung Austria, Gußhaustr. 8, 1040 Vienna, Austria

1 Introduction

The relevance of excellent science for the European Research Area (ERA) is undoubted. In this regard, the optimal use of the available human capital is crucial. Therefore, many funding programmes of the EU and the Member States aim to enable young researchers to follow their research ideas independently and to establish themselves within the respective science systems. In the career of a researcher, the post-doc phase is a decisive one. It is in this phase that the transition from an early stage researcher to an independent and established researcher takes place. In order to be recognised as a researcher in the national science systems, the post-doc researcher has to prove both excellent research performance and to show skills as an independent investigator. The latter aspect is to be demonstrated by raising external research funding, the supervision of Ph.D. candidates, the management and guidance of research groups and teaching experience, to name but the most common competences required. However, the opportunities to enable consolidating one's own research profile and position in the science system are rather few. As a consequence, many are forced to leave the science system for good during the post-doc phase in the absence of perspectives. In recent years, a general trend towards the prolongation of the post-doctoral phase can be observed, as a result of a cut-back of permanent research positions and an increase in short-term contracts. This situation is especially true for the Austrian research system.

The Austrian science system is a typical example of a "European continental model" (Huber et al. 2015). According to Huber et al. (2015, p. 21f.), the recruitment of professors is regulated by formal rules ("Berufungsverfahren" and in many cases the obligation to write a habilitation thesis). The system is characterised by few fixed term professorships (chair system) and the absence of tenure track options or fixed or at least long-term positions in addition to the professorship, such as lecturer positions or associate professor positions. The system is still ruled by seniority and informal agreements. Young researchers are very dependent on the chair holder under whom they work and the freedom for pursuing his/her own development is limited. The usual post-doc positions are limited with regards to contract duration to a maximum of 6 years, after which period the researcher cannot be employed anymore by a university in a non-permanent position.

Another trend observed in many European countries is the increase of programme- and project-funded research, in contrast to funding types that are directed to the university as such, e.g. basic and institutional funding. More and more funds are directly allocated to individual researchers or research groups based on competition. Universities encourage their researchers to apply for external funding e.g. from state agencies, the EU level or private entities (van Arensbergen and Pleur 2014). The group of post-docs especially has to rely increasingly on such external funding in order to be able to continue its research and develop the necessary competences to be appointed to a permanent position in the research system.

The Austrian START programme is designed to address the current situation of post-doctoral researchers and aims at closing the funding gap in the post-doc phase by creating a "protected space" (Huber et al. 2015) that enables the researcher to consolidate her/his position in the research system. The programme's main aims are, on the one hand, to support excellent and cutting-edge research and, on the other hand, to qualify the grantees for a permanent research position in the (Austrian) science system.

The aim of the programme evaluation was to inform about the effects the programme has generated since its start in 1996 and especially investigate whether the programme's

main objectives have been achieved. It was tendered as an "impact analysis" as not only the effects on the direct beneficiaries but also on group members, the host universities and the Austrian science system as such should be taken into account. Therefore, not only direct outputs and outcomes for the beneficiaries, but also broader impacts were at least qualitatively addressed. However, the evaluation as well as the discussion in this article focused on the direct beneficiaries (START grantees). With regards to different impact categories,[1] economic and wider societal impacts are only touched upon in the discussion of human capital development, especially related to the career skills of the grantees and the development of a pool of young researchers within the START project groups, but also the attractiveness of the Austrian research and innovation system.

For the purpose of this paper, the analysis focuses on the following two research questions:

- What effects has the START programme on the scientific performance of the START grantee over time, especially during and after the funding?
- What effects has the START funding on the career path and the pace of career development of the grantees?

These two research questions are investigated both with quantitative and qualitative methods whereas the analysis of the scientific outputs is based on a control group approach. This allows the quantification of impacts and attribution of them to the programme, while at the same time providing an explanation on the how and why of the intervention mechanisms of the START programme on different levels of the impact chain. The strengths of the analysis lie in the combination of approaches and research methods.

Next to the START grantees, further stakeholder groups (especially group members and host institutions, see Fig. 1) are affected by the START programme. With its outstanding position within the Austrian research funding landscape, it has furthermore also shaping effects on the Austrian research system, especially with regards to the strengthening of the capabilities of the Austrian science system. Furthermore, it affects the host organisations and the members of the research groups too. However, the analysis is focused especially on the level of the START grantees and partly on the group members and the host institutions. The analysis of the Austrian science landscape is less exhaustive, due to time and financial constraints of the evaluation. This also constrains this contribution, as the evaluation provides the data for this paper. The following impact pathway diagram has been the basis for the definition of outputs, outcomes and impacts.

The evaluation of the START programme was commissioned by the Austrian Science Fund (FWF). While its primary aim was to explain the impacts of the programme, specific questions of interest for the commissioning body with regards to accountability purposes of the past funding and the future development of the programme were not excluded from the analysis. However, they are not discussed in this article. The full evaluation study including recommendations for the future development of the programme is an open access publication and can be downloaded (Seus et al. 2016).

The article is structured as follows: Chapter 2 provides a short summary of existing literature in the field of evaluating impacts of research programmes for postdoctoral researchers. Chapter 3 describes the methodology chosen for the evaluation and provides a short overview of the START programme. Chapter 4 outlines the findings following different analysis levels. Chapter 5 discusses the limitation of the approach.

[1] As Roessner (2000) already showed, the terms inputs, outputs, outcomes, and impacts are not always used distinctively and without any overlaps.

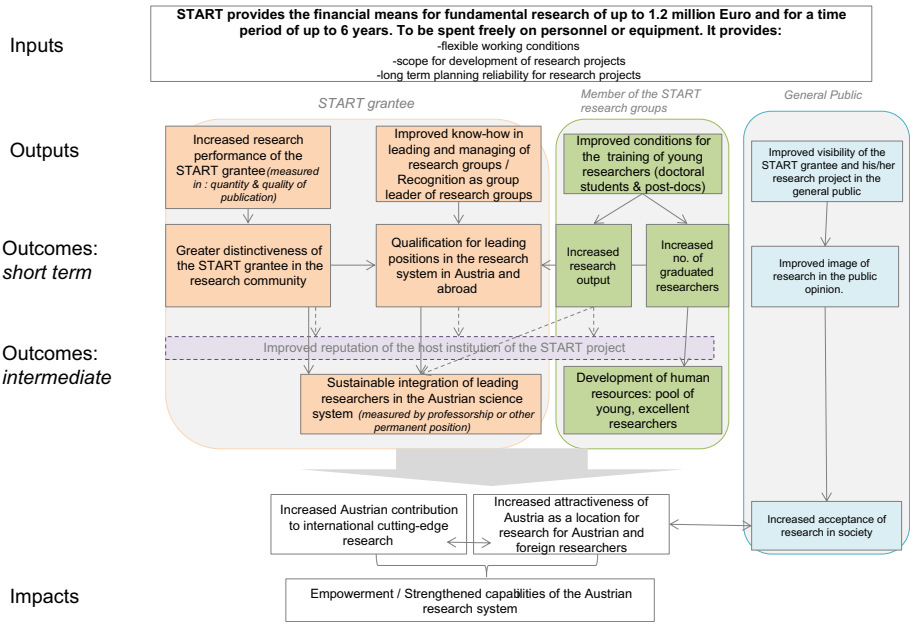

Fig. 1 Impact pathway diagram of the START programme (own compilation)

2 Literature on impact analysis of research funding programmes

In the European context, the START programme is not unique; several other countries have similar programmes for the target group of post-doc researchers (see detailed description of the programmes below). With a growing number of funding programmes especially targeting post-doc researchers, the number of studies on the effects of these programmes on scientific performance and career development has also increased. In most cases, these studies take the form of commissioned (evaluation) studies, which systematically investigate different types of impacts of a research programme on the beneficiaries and/or other stakeholder groups.

This is for example the case for the evaluation of the German Emmy Noether programme which is the German counterpart of the START programme (Böhmer et al. 2008) and a study comparing the Emmy Noether programme with similar research group funding programmes which exist in Germany (Hornbostel et al. 2009). In Sweden, a study on the Swedish Individual Grant for the Advancement of Research Leaders programme (INGVAR) investigated the scientific performance (publications), the international connectivity and the connected outputs such as patents or the reception of further funding. Career aspects such as positions held in the science system are however not examined. (Melin and Danell 2006). A study conducted by the Netherlands Bureau for Economic Policy Analysis investigates the career effects (sector of employment and type of contract held) of the "Innovation Research Incentive" (IRI) grant on the funded post-doc researchers (Gerritsen et al. 2013). All three studies evaluate the impact of the funding by using a regression discontinuity approach i.e. comparing the group of grantees with a group of applicants who were almost successful in obtaining the grant. This is a common practice in the field of research funding and especially with highly competitive grants in the late

post-doc stage, as in most of those programmes a two-step approach applies. This means that the best candidates are selected based on a proposal and then go through a second selection process (e.g. a hearing). The candidates just below the cut-off line are therefore almost equally good candidates and have very similar features with regards to scientific performance and skills to the candidates selected. The German and the Swedish study base their analysis on online surveys with both groups, as well as bibliometric analysis of the scientific performance. While the main analysis of the Swedish INGVAR programme is based on quantitative data, notably bibliometric analysis, the evaluation of the Emmy Noether programme also includes a qualitative interview part. Furthermore, it compares the Noether programme to similar existing junior group leader funding programmes in Germany by sending online surveys to grantees and applicants of four other programmes. The study of the Dutch IRI programme is a purely quantitative and econometric analysis using data from different public databases that give indications on the current job status of the grantees and the control group. Here, only a regression analysis is performed. At European level, evaluations of funding programmes dedicated to supporting young researchers are conducted regularly and made available to the public. This is the case for the European Marie Curie Programme which in its most recent evaluation assesses the impacts of the funding programme on research careers (European Commission 2014a) or a broader study on the contributions of the European framework programmes to the development of the human capacities at the level of the individual researcher (skills and expertise of researchers, career effects including contractual effects), the research group level and the system level (European Commission 2014b). Closely related to the career effects is the international mobility of researchers (not exclusively focusing on the post-doc phase) and the tendency to seek a position in a foreign country which fits under the discussion of brain drain (Conchi et al. 2014; Böhmer and Hornbostel 2009; Meyer and Bührer 2014; PREST 2002; European Commission 2014a, b).[2]

One of the most elaborated impact analysis in the field of individual funding programmes is the EURECIA project "Understanding and Assessing the Impact and Outcomes of the ERC and its Funding Schemes", investigating the impacts of the ERC starting grant scheme, funded by the European Research Council (Nedeva et al. 2012). The aim of this project was, most of all, to develop and test a conceptual framework for the impact analysis of research funding schemes. A test phase for the methodology had been included, which provided some insights on the impacts due to the funding programme. Impact categories that were covered by the EURECIA project were: the impact on the researchers' careers, on the research content, on the host institutions, on the national funding landscapes of EU Member States and the European funding landscape. For this 19 matched pairs of funded researchers and control group researchers were investigated in depth. Each impact category uses a different methodological approach. The ERC starting grant has been monitored since 2009 in a more traditional manner. The 2015 evaluation report focuses on the effects of the grant on the grantees in terms of research independence, development of skills and competences and career development, among other aspects (Huber et al. 2015). All these studies of European programmes use a control group approach, comparing funded researcher to non-funded researchers. In the case of both studies on ERC, the

[2] The concept of "brain drain" refers to scientists who migrate from one country to another with no intention of returning (Grubel 1994)—an action that has international, economic and political impacts, especially in developing countries (see e.g. Lowell 2002). The brain drain approach argues that countries lose human capital if scientists go overseas to study or work, as they might decide to remain there. Thus, the emigration of highly skilled scientists results in a human capital loss ("brain drain") for the former home country and human capital earning ("brain gain") for the respective host country.

control group consists of unsuccessful candidates, while for the studies on Marie Curie fellowships and the FP, a control group is made up of non-funded researchers with similar characteristics to the funded persons and is generated via a bibliometric database. All studies use a more or less sophisticated traditional triangulation approach consisting of an online survey, a bibliometric analysis of the scientific performance and qualitative interviews.

Regarding the scientific performance, the study on the Swedish INGVAR programme and the Marie Curie fellowship found a slightly increased citation rate of funded researchers, while there is no difference compared to the treatment group with regard to the scientific output. The evaluation of the Emmy Noether programme found no evidence that the grantees publish more or in journals cited more often. Regarding the career development in the science system, all studies, except the one investigating FP researchers, see a positive effect of the funding, which translates into a higher rate of appointments to full professorships of the funded researchers. However, the incomes of the funded researchers are not higher than the ones of the control groups; also the number of researchers with a permanent contract is not higher in the group of grantees.

Current discussions in the field of science policy evaluations are moving away from the pure measuring of outputs as e.g. done with bibliometric methods and explore methods and approaches to catch broader structural effects of science policy and programme funding. Recent publications have tried to develop frameworks and methods to clarify and assess social or societal impacts of research (Morton 2015; Joly et al. 2015; Spaapen and van Drooge 2011). The challenges of assessing broader outcomes and impacts of research policies on indirect beneficiaries and a wider public are recurrent in all debates: the problem of attributing the effect to a specific policy measure, the interference with other funding or contextual events and the sometimes considerable time lag between the programme's implementation and the occurrence of impacts (Penfield et al. 2014). Several authors therefore suggest using the concept of contribution instead of attribution (Mayne 2001; Morton 2015; Nedeva et al. 2012) and a technique now commonly used in evaluation studies is a mixed methods approach, which combines quantitative and qualitative data and data from different sources.

The present contribution adheres to this direction of thought and takes the limitations and considerations into account while exploring structural effects of the START programme at the different levels of the Austrian national science system.

3 Study design: methods and data sources used

3.1 The START programme

The START programme provides outstanding young researchers of any discipline with a considerable amount of funding (up to 1.2 million Euro) that can be freely spent on personnel and infrastructure and provides the financial security to carry out his or her research project for a time period of up to 6 years. It is one of the most highly supported research programmes and the most prestigious individual research grant for postdoctoral researchers in Austria. The primary objectives of the START programme are twofold: (1) promoting excellent science at the international top level; (2) support young researchers in gaining a permanent position in the Austrian or international science system. The sole selection criteria being scientific excellence, the grant has been awarded to researchers

from the whole spectrum of research disciplines, covering natural sciences, life sciences, social sciences as well as humanities.

The programme has been managed by the Austrian Science Fund (FWF) since its creation in 1996. Between 1996 and 2014, a total of 114 START grantees were funded, amounting to a total funding sum of approx. 32.6 million Euros. Taking into account this considerable amount of funding spent on basic research and on researchers, the programme as such is expected to strengthen the capabilities of the Austrian science system overall. Hence, one of the preconditions of the funding is that the grant holder is affiliated with an Austrian research institution where the majority of the research project has to take place. As a flagship programme for talents from all over the world, the START programme addresses researchers based in an Austrian institution as well as foreign researchers equally. Although the START programme is not primarily designed as an instrument to incentivise Austrian researchers located in foreign research institutions to return to Austria or as an instrument to attract foreign researchers to the Austrian science system, it can be used for that purpose and has been so in the past. Besides the direct and primary beneficiaries, the researchers, the START programme has effects also on two other categories of beneficiaries: namely the START group members and the host institutions of the START grantees.

3.2 The evaluation design: combining quantitative and qualitative approaches

As the evaluation was intended as an impact evaluation, the study design should, on the one hand, quantify the results as far as possible and, on the other hand, give an indication whether those documented effects could be attributed to the START funding. Consequently, the evaluation design chosen was a quasi-experimental one, which incorporated a counterfactual element in form of a control group. A bibliometric analysis as well as an online survey was used to compare the group of START grantees to its control group.

The strength of experimental designs is that they assess whether measured effects can be attributed to a programme or not. However, such a design does not explain how and why policy interventions work (or do not work) (Mayne 2012, 2001; Morton 2015). The opening of this black box is however crucial for an in-depth understanding of the programme, and also to provide meaningful policy recommendations. In order to counterbalance the weaknesses of traditional quantitative impact analysis (Leeuw and Vaessen 2009), the evaluation was complemented by the analysis of further data, using different data collection methods and sources (mixed methods approach). With the aim of triangulating the evidence, the analysis combined different data sources as well as qualitative and quantitative analytical methods, wherever possible and deemed meaningful.

The following methods and sources were selected for the evaluation: a survey of a comparison group that consisted of unsuccessful applicants and qualitative analysis in form of case studies which included evidence from interviews, a documentary review, the analysis of existing monitoring data and a workshop with experts from the Austrian science system. In the following, the different methods used are described in more detail.

3.3 The control group

The control group was created by using the Scopus. In a first step the START grantees were identified in the Scopus database. Scopus has the great advantage that it provides unique identification codes for each author. Once the researcher has been identified through

his/her (or sometimes several) ID(s), the publication output can be traced. The unique ID(s) for each researcher have the advantage that duplications can be excluded. With the exception of two START grantees, all 114 grantees could be identified in the database. The two grantees that could not be identified belong to the group of social scientists and humanities (in these particular cases: law and history). The coverage of the social sciences and the arts and humanities in bibliometric databases is generally considered to be rather low. There are various reasons for this, such as the language bias against non-English language publications and different citation and publication behaviours in these scientific fields (Chi 7/15/2013; Norris and Oppenheim 2007; Archambault et al. 2006). In a second step and based on the information that could be obtained for the START grantees, a randomly selected control group was created that consisted of researchers with comparable characteristics at the time of the awarding of START; i.e. a person that could have potentially also been a START applicant. Thus, the control group consisted of a "twin" for each START grantee that had been generated through a propensity score matching process. The one-to-one search for a twin allows for a thorough manual check of the selected twin as being suitable for the analysis. The following characteristics were used:

- the research field in which most publications were published;
- the gender of the researcher, which was automatically extracted from first names;[3]
- the year of the first publication in the Scopus database; the goal of this selection criterion was to have a control group with a similar distribution of career statuses as in the START grantees set. Since Scopus does not include information such as research position or degree, the number of years of scientific activity was calculated.
- at least 5 publications in 5 different years with an affiliation with an Austrian institution in order to be considered "Austrian"; e.g. a person with personal involvement in the Austrian science system and therefore sufficient knowledge and interest in the START programme;
- a similar number of publications up to the beginning of the START funding, in order to include only researchers with a similarly high scientific performance and therefore a potential candidate for the START programme. Here, ranges for matching have been defined, instead of looking for the exact same number of publications.
- a similar citation rate up to the beginning of the START funding, in order to include twins with a similarly high performance. Here too, ranges had been defined.

In order to guarantee a sufficiently high response rate to the online survey, up to *three* "twins" (if available) for each START awardee were randomly selected as control group twins. For the bibliometric analysis, only one twin, here called the *'main twin'* out of the three was selected. This was the twin that had the closest matching characteristics to the START grantee. This selection was then cross-checked manually.

For the process of twin generation compromises had to be made in some cases. For example, if no twin could be found when applying all criteria mentioned above, the criteria "research field" (defined as the field with most publications) was interpreted more widely and those researchers were included that had at least one publication in the main field of the START awardee. This was seen as a legitimate approach, as researchers often publish across several disciplines. In those cases, in which no twin could be generated, the characteristic "gender" was dropped. As this was only the case for some cases, no bias is

[3] Fraunhofer ISI has implemented an Oracle-SQL version of this database and systematically added further data and information to the database. Among the extensions are regionalisation (NUTS1, NUTS2, and NUTS3) of the EU-27 Member States or the definition of the researcher's sex via the first name.

expected. In total, 112 START grantees have been compared to 108 twins, as two START grantees could not be identified via Scopus and for another two START grantees no twins could be found.

The use of such a control group is rarely applied in similar studies. Often, the comparison group used consists of unsuccessful applicants (e.g. Huber et al. 2015; Böhmer et al. 2008). The use of unsuccessful applicants for the comparison relies on the assumption that unsuccessful applicants are at the time of grant allocation almost equally qualified (e.g. in terms of scientific performance and future research potential) to the researchers who finally received the grant. However, this approach implies that other variables, such as gender and discipline are not controlled. Furthermore, the use of unsuccessful applicants introduces a bias, as the analysis is limited to the perspective of individuals who know the programme or have a certain proximity to the respective funding bodies.

Our approach is able to cope with a range of potential sample selection biases and allows for the attainment of comparable and balanced samples of researchers. The START grantees and the control group should only differ by their treatment but not by other structural variables, like the specific field.

3.4 The comparison group

As in other studies, the evaluation of the START programme could also make use of a comparison group, consisting of unsuccessful applicants. Since 2006 the application process for the START programme has been a two-step process, comprising a written application and a peer review and in a second selection step a hearing in front of the START jury. The applicants who are invited to the hearing have already passed a rigorous selection process in which their scientific performance has been ranked as one of the best in their disciplines. As discussed above, the unsuccessful applicants are assumed to have a similar scientific performance and profile to the START grantees, as the final selection between two or several candidates is in those cases based on minimal differences. Due to their knowledge of the START programme and their experience with the FWF as a funding body, they provided valuable information on the START programme and its impact mechanisms as well as providing valuable explanatory information for the evaluation.

Other than for the control group, these unsuccessful applicants were only invited to participate in an online survey.[4] No comparison of their scientific performance with a bibliometric analysis took place. Since 2006, there have been 57 applications; amounting to 49 individuals (due to double applications or successful START applications in the following years).

3.5 Bibliometric analysis

In order to assess the scientific performance of the START grantees, and especially to show changes of the publication performance over time, a bibliometric analysis was performed, comparing the group of START grantees to the control group. The scientific performance was calculated using five indicators as proxy, namely:

- number of publications,

[4] It was not possible to conduct a meaningful bibliometric analysis on the comparison group as the group of people was too small ($n = 49$) and time spans too restricted to be able to conduct a before—during—and after analysis for this group.

- citation rate,[5]
- number of publications written with co-authors,
- number of organisations per co-publication,
- number of countries of origin of the co-authors.

In order to describe the change in the scientific performance over time and allow statements on the immediate and medium-term impacts of the START funding, the above mentioned indicators were measured for three different timeframes, corresponding to different periods in the research life of the START grantee. The indicators for the control group twin were calculated following the specific timeframe of its corresponding START grantee.

In order to take into account the time lag between the generation of research outputs and their publications in scientific journals, the period of analysis did not correspond completely to the funding period of the START programme, but was adjusted for the time lag.

The three periods are defined as follows, each of them including on average 5 years:

- *Before* the START funding period: this period is defined as 3 years before the beginning of the START funding until after the second year of funding. *(n = 103 START grantees; 102 control group)**[6]
- Period *during* which START funding was received: this period starts in the third year of the START funding until 1 year after the end of the funding period. *(n = 51 START grantees; 54 control group)**
- *After* the START funding period: this is defined as the period from 2 to 7 years after having received the START funding. *(n = 38 START grantees; 36 control group)**

Besides the measurement of publication output, the bibliometric database Scopus was also used to identify whether the group of START grantees differs from all Austrian researchers with regards to the three characteristics: gender, geographical location of the affiliated research institution and discipline. Austrian researchers are defined as those researchers who had for at least 3 years published as a researcher affiliated with an Austrian research institution.

3.6 Online surveys

The online survey aimed to quantify the effects of the START grant on career aspects, especially the sector and location of employment, the position held and the pace of career development.

Three different target groups were surveyed in the course of this evaluation:

- *The START grantees* based on the FWF database, all START grantees who had received a grant between 1996 and 2014 were asked to participate in the survey. Overall these were 114 individuals. 95 of them participated in the online survey.

[5] As the selection of the control group also controlled for the discipline, the analysis used the citation rate of each researcher and not the field-weighted ones. The analysis of the citation rate is an average of all citation rates of the START group and the control group. It therefore gives only an indication of changes, but not of productivity in individual disciplines.

[6] * The differences in the number of grantees and control group researchers included in the sample is due to the fact that it was not possible to generate a twin for each START grantee or that data were missing. The drop of persons between the three periods of analysis is due to the longitudinal design of the analysis and the fact that not all grantees have finished the funding period yet.

- *The control group* (labelled "CG" in Fig. 5): For each START grantee three "twins" were generated and randomly selected from the Scopus database. In total 307 researchers could be identified by matching the criteria for a control group. 27 individuals could not be reached, mainly due to invalid email addresses. In total, 75 researchers completed the online survey.
- *A comparison group* (labelled "CS" in Fig. 5): This group consisted of unsuccessful applicants to the START programme. In total, the questionnaire was sent to 49 individuals of whom 25 completed the online survey.

The surveys took place between February and May 2014. In the following tables the distribution between the three surveyed groups with regards to gender, disciplines and scientific age is shown (Tables 1, 2 and 3).

3.7 Qualitative analysis

In order to obtain a deeper understanding of the cause and effect mechanisms of the funding, the results of the control group analysis were complemented by further evidence of a mostly qualitative, but also of a quantitative nature: Eight case studies of START projects were conducted that combined the views of START grantees, representatives of the host institutions of START projects, and START project group members (23 interviews in total). Interviews with institutional stakeholders (FWF, Federal Ministry of Science, Research and Economy) and the START jury (8 interviews in total) and analysis of programme documents, project reports, and monitoring data complemented the evidence base of the evaluation. The qualitative analysis took place between May and July 2014. The themes discussed in the interviews were the following: possibilities of the START programme especially for conducting research in comparison to other funding opportunities; meaning of the START grant for the scientific career (in Austria); effects of the START programme on group members (research opportunities and career prospects); role of the host institution; general effects of START on the Austrian scientific landscape; strengths and weaknesses of the START programme; the motivations for application; assessment of administrative requirements and the programme management. An expert workshop, gathering institutional stakeholders from the BMWFW and the FWF, representatives from Austrian research organisations and START grantees was held on 24th September 2015 to discuss the preliminary findings of the evaluation.

Table 1 Gender distribution in the samples

	Survey responses		Overall sample (persons to whom survey has been sent)	
	Male (%)	Female[a] (%)	Male (%)	Female (%)
START grantees	83.3	16.7	84.2	15.8
Control group (CG)	80.0	20.0	85.6	14.4
Comparison group (CS)	72.0	28.0	75.5	24.5

[a] The higher response rate of women in surveys is a recurrent feature and has been discussed in the literature (e.g. Bogner and Landrock 2015; Preisendörfer and Wolter 2014; Krumpel 2013)

Table 2 Disciplinary distribution of survey respondents in %

	Survey responses		
	START grantees	Control group (CG)	Comparison group (CS)
Physics	27.4	13.3	11.5
Biotechnology/medicine	23.2	20.0	38.5
Natural sciences	21.1	34.7	3.8.
Mathematics	15.8	21.3	19.,2
Social sciences and humanities	12.6	10.7	23.1
Others	0.0	0.0	3.8

Table 3 Scientific age of survey respondents, measured as difference between 2015 and year of Ph.D. award (in years)

	Survey responses		
	START grantees	Control group (CG)	Comparison group (CS)*
5 years or less	0.0	7.0	4.0
6–10 years	21.5	21.1	32.0
11–15 years	35.5	22.5	52.0
16–20 years	17.2	19.7	12.0
21 years and more	25.8	29.6	0.0
N =	*93*	*71*	*25*

* Researchers from the comparison group have only be included since 2006; consequently they have not yet reached a very high publication age, compared to the 2 other groups

4 Findings

4.1 Micro level

At the micro level, two different impact categories were investigated. They target the effects of the programme on the funded grantee itself; firstly concerning changes in the scientific performance during and after the START project funding and then, secondly, the effects on the career development of the grantee.

4.1.1 Scientific performance

The impacts of the programme on the scientific performance of the START grantees were analysed using a bibliometric analysis of publications from the group of START grantees and from the control group (CG). The results of the bibliometric analysis were then triangulated with evidence from the surveys and the case study thus enabling pushing the interpretation of the results one step further.

4.1.1.1 Development of scientific performance over time The result of the analysis shows that the START grantees increase their performance on almost all indicators during the three periods investigated. Along the three periods, the START grantees publish more in absolute numbers, are increasingly active in publishing together with other researchers and with other research institutions and are also well connected to the international research community with an increasing number of publications written with institutions outside Austria (see Figs. 2 and 3).

At first glance, the citation rate shows a slightly different trend than the other indicators. It would, however, be premature to conclude from the decreasing citation rate that there is a decrease in performance. Here scaling effects might enter into play: When the amount of publications increases, but the amount of highly-quoted publications remains constant (which is often the case), the average citation rate decreases.

In another reading, the downshift of the citation rate during and after the START funding could also be a direct effect of the START funding: namely the result of a newly gained scope for testing new and unconventional research fields. Evidence from the survey and the case studies lead to the assumption that some START grantees ventured into a re-orientation of their research towards new or more unconventional research fields for which he/she is not (yet) visible in the research community. In the interviews with the START grantees it was highlighted that the duration and the amount of funding provides a certain safety and allows taking one's time. This argumentation is supported by the survey of START grantees, in which 75% of the respondents state that the START programme had been conducive to accessing new fields of research (see Fig. 4). The slight increase of the citation rate in the period after the START funding can be seen as an indicator for a promising outlook with regards to the future scientific performance of the former START grantees, once they have settled in their new research fields. Earlier studies on citation rates of FWF-funded projects and programmes come to a similar conclusion (FWF 2007).

Furthermore, the citation rate allows conclusions on the selection process: The high citation rate of 11.8 before the START funding indicates that researchers with an outstanding scientific performance were selected. The goal of the START programme to fund outstanding researchers in Austria can be seen as having been achieved.

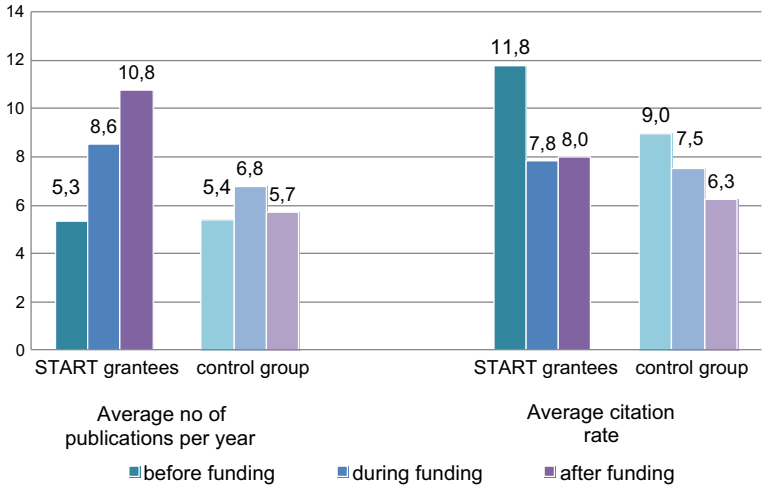

Fig. 2 Average number of publications per year and average citation rate

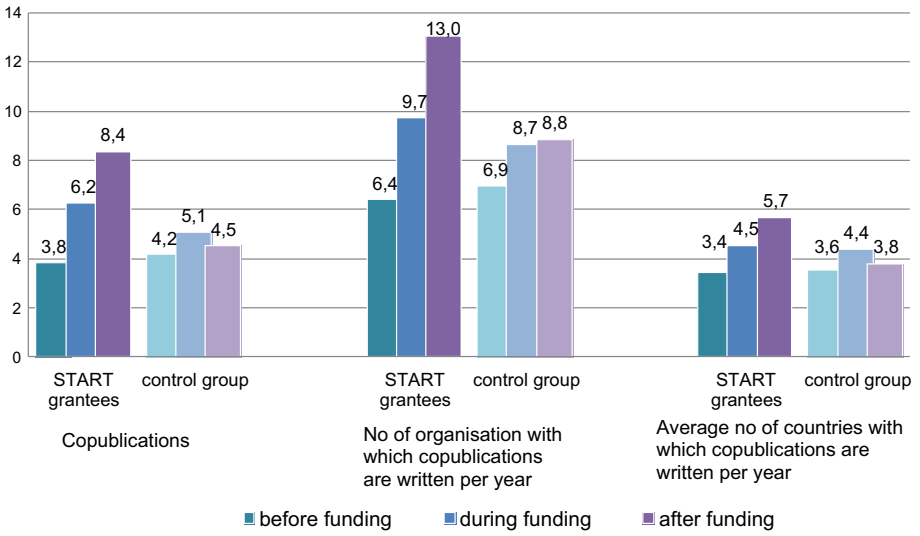

Fig. 3 Co-publications, collaborating organisations and countries

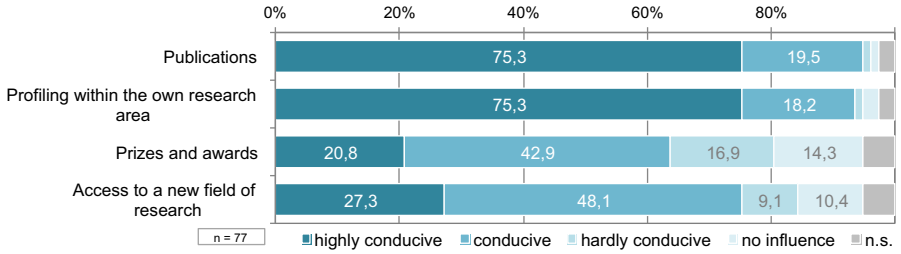

Fig. 4 Effects of the START project on the performance in the area of research (START grantees). Question asked: "How has the START project affected your performance in the following areas?" / n.s. = no statement

4.1.1.2 Scientific performance of the START grantees compared to the control group The comparison with the control group shows that the START grantees perform better throughout for all indicators and also during the different periods analysed (see Figs. 2 and 3). The differences between the groups are statistically significant for all indicators (calculated for the period "after funding"), except for the citation rate.[7] The fact that the performance of the control group decreases for all indicators except the cooperation with organisations for the period "after the funding", while the START grantees' performance increases further, suggests that the START programme contributes to boosting the scientific performance of the grantees, especially in the medium run. The findings from the survey of START grantees support the argumentation that the START programme contributed to strengthening the grantees' scientific performance. Almost all survey respondents state that the START programme was conducive for their publication output and establishing their own profile in the respective research area. More than 60% indicate that the START programme was conducive to receiving further (academic) awards (see Fig. 4).

[7] The Mann–Whitney U Test has been used to calculate the similarities between the two groups.

When asking the comparison group, i.e. the candidates to the START programme, whether the missing START grant had impacted on their scientific performance, 70% of the respondents stated that it had hindered their publication performance; almost 60% state that the START programme would have been beneficial for establishing their own profile within their own research area, and 50% state that it had affected the access to new fields of research negatively.

We concluded from the different analysis that the START programme has had a considerable positive impact on the scientific performance of the researchers receiving START funding during the funding but even more, for their future scientific performance. Due to the control group approach we could furthermore conclude that the changes in performance can be attributed to a considerable extent to the START programme.

4.1.2 Career effects of the START programme

In order to assess the contribution of the START programme on the career of the grantees, a survey of the START grantees and of the control group (CG) as well as the comparison group (CS) was used, complemented by interviews of the START grantees. The following three indicators were analysed in order to assess the impacts of the START programme: the employment sector of both groups after the end of the START project; the position held by the START grantees in the post-START period and the "pace of career development" measured as number of years between the year of being awarded a PhD and the year of the first full professorship position. Other than for the scientific performance, the impacts of the START programme on the career of the grantees cannot be assessed as positively.

4.1.2.1 Employment sectors The study results show that *all* START grantees who responded to the survey and had (already) finalised their START project, remain in the research system. Around 64% of START grantees remain in a scientific institution in Austria. The majority of former START grantees remain in the institution in which he/she was employed for the START project (57.8%), while 35.6% have chosen to pursue their career in a research institution abroad. Less than 5% have chosen an Austrian research institution different from their START project institution. However, the employment patterns of the two control groups are rather similar and show no statistical difference to the group of START grantees.[8] Just like the START grantees, most survey respondents from CG and CS today work as researchers in a research institution, either in Austria or abroad.

For the researchers who left Austria, the reasons for joining a foreign research institution are quite similar in all three surveyed groups. The main motivation is an attractive job offer from abroad. The second main reason for the START grantees is the lack of an appropriate position in Austria. Furthermore, both the control and comparison group also state "better long-term career prospects" and "better conditions for researchers" as a motive for leaving Austria. These last two aspects are quoted less frequently by the START grantees and are a hint that the START project provides a good starting position to establish oneself in the Austrian science system.

These arguments are also put forward in the case studies. The reception of the START grant is often seen as the decisive factor in the decision to remain in Austria. Two interviewees report that without START they would have left the country, one returned from the USA because of START.

[8] Fisher's Test, significance level $p < 0.05$.

4.1.2.2 Positions held by the START grantees Regarding the nature of positions[9] held in the science system, START grantees do not differ from the two non-funded groups. The control group and the candidates for the START programme show a similar pattern with regards to job positions (Fig. 5). 80% of the START grantees responding to the survey hold a permanent working contract after the completion of the START programme. In comparison, only 64% of the surveyed START grantees held a permanent contract during the START project. Before the START programme, this was the case for only 28% (see Fig. 6). In recent years (since 2006), it seems that the reception of the START grant goes along with a permanent position for more than half of the START grantees. In the interviews, the younger START grantees with a project at universities report that they have gained temporary professorship positions with a view to gaining a permanent position after START.

4.1.2.3 Pace of the career development On average, START grantees get appointed a full professorship position 7.5 years after obtaining a PhD. However, for this indicator no statistically significant difference can be found between the group of START grantees and the control group either. START grantees do not get appointed to a professorship earlier than the control group or the comparison group.[10] Around 80% of all surveyed researchers are appointed as professors between their fifth and the 15th year after receiving their doctorate. In the control group this is the case for 78% and for 82% in the comparison group.

Since 2006, START researchers as well as researchers from the control group have been appointed earlier to a professorship position than researchers and their "twins" who have received the funding before 2006.[11] This might be related to changes in the Austrian research system, such as the introduction of tenure track positions.

Although from a time perspective the START programme does not speed up the research career, the START programme is perceived as an impetus for the research career development by the START grantees. Almost all of the responding START grantees (97.4%) are of the opinion that START has strengthened their career prospects. Furthermore, 60% of the responding START grantees are convinced that they would not have reached their current positions without the START programme. This is also confirmed in the case studies. However, in the comparison group only one-third of the respondents think that the participation in the START programme would have permitted them to reach their current positions earlier.

4.2 Meso level

At the meso level two types of impacts were analysed. First the impacts on the group members were investigated. A second indicator examines the effects the programme had on the research institutions hosting the START group.

[9] For this study three different professorships have been distinguished, based on the former and recent Austrian research system: full professorship (Universitäts- or FH-ProfessorIn); Associate Professor (Assozierter(e) ProfessorIn, ehem. DozentIn); Assistant Professor (AssistenzprofessorIn).

[10] However, this result has to be interpreted with caution, as it is based on a relatively small number of survey respondents for all three groups: n(start) = 64; n(control group) = 41; n(comparison group) = 17.

[11] Pearson correlation index for the START grantees is −0.705; for the CG −0.556); no data for the comparison group is available, as data for this group exists only from 2007 onward.

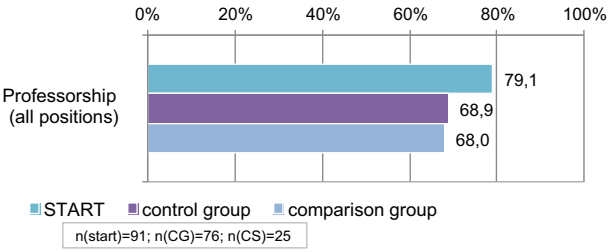

Fig. 5 Highest position in the research system reached to date

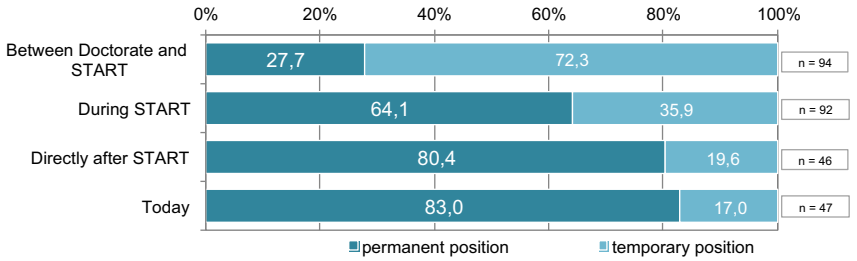

Fig. 6 Permanent versus temporary work contract (START grantees). Question asked: "Which type of work contract did you hold during the following career phases?"

4.2.1 Impacts on START group members

As a direct consequence of the creation of START research groups, young researchers get the possibility to participate in cutting-edge research. The monitoring data of the end-of-project reports hint at the number of young researchers who were supervised by START grantees at different levels of their education. The following table shows the number of theses supervised by 53 START grantees who had finalised their projects by April 2015 (Table 4).

All START grantees except one (out of 51) had supervised at least one doctoral thesis. Furthermore, 42 out of 51 START grantees had supervised a first degree thesis at least once, such as a bachelor's, master's or diploma thesis. 30 out of 51 START grantees had supervised one or more habilitation theses.

Table 4 Total number and average number of supervised theses in START research groups

	Total no. (completed and ongoing*)	Average (per START grantee)	Completed theses	Average (of completed theses)
Diploma/master's/ bachelor's theses	272	5.33	240	4.52
PhD/doctorate theses	298	5.62	222	4.19
Habilitation theses	75	1.47	32	0.60

Total number of START grantees included in the computation: 53

* Ongoing = not finalised by the end of the START project

NB: The figures are taken from the end-of-project reports of the START grantees and were provided by each grantee on her/his own calculation or estimation. The figures must therefore be interpreted with a certain caution, also taking into account that the START grantee looks back on a long time period of up to 7 years

From the perspective of the START grantees and the group members interviewed, being part of a START project group provided young researchers with the opportunity for participating in cutting-edge research within an international environment and with good contacts to their research community and a well-funded project budget (e.g. for conferences, workshops or visiting scholars). Furthermore the START project offered the possibility to receive a contract with a relatively long duration, as it offers financial security and job stability for up to 6 years. In this context the flexibility of the programme is important: e.g. a PhD contract can be transformed into a post-doc researcher contract without bureaucratic hurdles. Both START grantees and research group members were of the opinion that the participation in a START group also seemed to strengthen the personal career prospects of the group members (see Fig. 7).

4.2.2 Impacts on institutions of higher education

The problem of correctly attributing the effects observed becomes increasingly more relevant when moving upwards in the impact chain (Nedeva et al. 2012). Establishing direct causal links to the programmes gets increasingly more difficult. However, as the START programme has been on the go for almost 20 years, it has been slowly integrated as a prestigious funding source for host institutions.

In the case studies it became obvious that the way of handling the START programme and the START grantees has changed over time in the Austrian research institutions. The "older" START grantees seemingly were left alone in preparing the START project, as well as during the project implementation phase. In recent years, institutional support has been provided to a greater extent by the host institutions. Also, host institutions have been more proactively approaching researchers whom they assess as being qualified for such a competitive programme in recent years. This has been done via the rectorate, the deanery or the service in charge of research promotion ("Forschungsservice"). Furthermore, excellent researchers from abroad who show interest in a position in Austria have been directed towards the programme. One interviewee in this context noted that the institution tried to look for projects which suit the research focus and the strategic development of the institution especially well. In the application phase, host institutions today offer support and consulting, such as feedback on the work plan and financial outlines, and workshops for applicants with current START grantees. Another more recent trend is the discussion of resources needed for the projects and the framework conditions beforehand. Also, the career perspectives after the end of the START projects are discussed. Today, most universities offer the START grantee a tenure track position and negotiate a qualification agreement ("Qualifikationsvereinbarung"). If the habilitation is completed within the period of the START project, he/she becomes an associated professor (permanent position). For those START grantees without the tenure track option, other options are created. All in all the negotiation power of the START grantee towards his/her host institution has

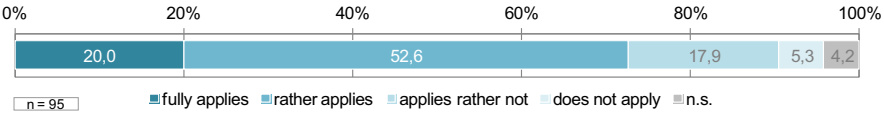

Fig. 7 Career prospects of the START group members (START grantees). Question asked: "To what extent does the following statement apply: The START grant strengthens the personal career prospects of the employees of the funded projects"

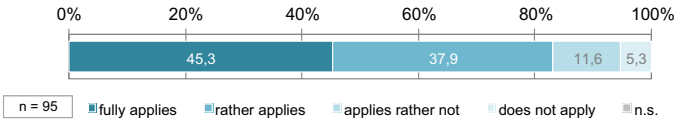

Fig. 8 Effect of the START programme on the negotiation position towards the host institution (START). (1) Question asked: "The START programme strengthens the negotiating position of the START grantees towards their research institution regarding the shaping of the research conditions."

increased considerably in the last years. This increased negotiation position was also highlighted by the START grantees in the survey and in the case studies (see Fig. 8). This holds true although the START programme, unlike similar programmes such as for example the ERC starting grant, does not include a formal negotiation procedure between the grantee and the host institution at the beginning of the project.

4.3 Macro level: Science landscape Austria

The impact of the START programme on the Austrian scientific system was investigated through two indicators: first, the distribution of START grants between certain disciplines and in certain regions (as a proxy for research institutions) was examined. This gives an indication of the concentration of scientific excellence or "internationally known hotspot of science". Second, the success rate for an ERC grant was taken as an indication for increasing overall scientific performance of Austrian researchers and showed an internationally strengthened Austrian science system.

4.3.1 Hotspots of scientific excellence

In the following sections we investigated whether the START programme had been distributed equally between the different regions, disciplines and between male and female applicants, as the sole selection criterion for the START grant should be scientific excellence. In order to do this, the distribution of the three characteristics mentioned within the START applicant and the START grantees group were tested against the group of all Austrian researchers (defined as researchers with at least three publications in which the researcher has been affiliated with an Austrian research institution). The grouping of disciplines is based on journal publications. There are certainly candidates for the START programme who research at the border of more than one of these five disciplines. As this, however, holds true for all three groups, the analysis can be judged as being adequate. To create the group of all Austrian researchers, the Scopus database was again used.

The analysis shows differences in distribution regarding regions and disciplines of applications compared to all Austrian researchers. This already is the case at the application stage. In terms of regional and disciplinary distribution of the START grantees compared to all Austrian researchers, it can be shown that there were certain areas (geographical and disciplinary) which received more than their proportional degree of funding. This means that researchers from particular regions (see Table 5) or from particular disciplines (see Table 6) did not apply or applied less frequently for START grants.

In particular, the share of applicants from the fields of mathematics and physics was significantly higher compared to the group of all Austrian researchers. Contrasting this, there were proportionally fewer applicants from the fields of biotechnology/medicine and natural sciences compared to the group of Austrian researchers.

Table 5 Regional distribution of START candidates and grantees

	R1 (%)	R2 (%)	R3 (%)
All Austrian researchers	54.46	19.55	25.99
Candidates for the START programme since 1996	71.8	8.89	19.82
START grantees since 1996	64.91	9.65	25.44

Region 1 (R1): Eastern Austria, including the Vienna region, Burgenland and Lower Austria

Region 2 (R2): Southern Austria, including Carinthia and Styria

Region 3 (R3): Western Austria, including Upper Austria, Salzburg, Tyrol and Vorarlberg

All three regions are significant at the level of $p < 0.01$; Mann–Whitney-U Test

Table 6 Disciplinary distribution of START candidates and grantees

	Biotechnology and medicine (%)	Mathematics (%)	Natural sciences (%)	Physics (%)	SSH (%)
All Austrian researchers	53.08	2.60	32.98	6.69	~ *30*[a]
Candidates	31.73	8.38	25.25	21.45	13.07
START grantees	22.81	15.79	20.18	27.19	14.04

[a] As Scopus represents the SSH (social science and humanities) disciplines only to a very limited extent (around 4% of all publications in the case of Austria), the She-figure report 2012 was used. It provides data for 2009: according to this source around 30% of Austrian researchers were active in the field of SSH; around 17% in social sciences and 14% in humanities (see pages 46 and 62). http://ec.europa.eu/research/science-society/document_library/pdf_06/she-figures-2012_en.pdf. The distribution of START grantees was the following: 5.3% from social sciences and almost 7.9% from humanities

In two studies of FWF-financed programmes, (Reckling and Fischer 2010; FWF 2007) the authors argue that Austria has a handful of high-performing and world-class leading disciplines (especially mathematics, physics and biology). We conclude that the concentration of START groups in exactly these disciplines and well-known research institutions, indicate the interpretation that nowadays there are "hotspots of research excellence", which to a certain extent reproduce themselves with regards to receiving START funding. This is e.g. the case for the physics department of the University of Innsbruck which has hosted two START groups and one Wittgenstein awardee[12] since 1996. But this concentration of START grants can also be observed in certain social science disciplines (specifically history and archaeology at the University of Vienna/Austrian Academy of Science). The results thus confirm the well-known Matthew effect (success breeds success) (Merton 1973).

[12] The Wittgenstein award provides recognition and support to excellent scientists who have already produced exceptional scientific work and who occupy a prominent place in the international community and have a permanent position as professor in one of Austria's research organisation. It is the most generously supported research award in Austria.

4.3.2 *Internationally strengthened Austrian research system*

Another finding of the evaluation was that the START programme contributed to a high number of Austrian researchers participating successfully in the European ERC starting grant programme, compared to other European countries. The obligation of a START applicant to also apply for the ERC grants in parallel to the START grant is certainly one factor supporting the participation in ERC grants.[13] Our analysis shows that the double application obligation contributed to this high rate of ERC projects carried out within Austrian research institutions (between 2007 and 2014, Austrian organisations have hosted 75 ERC grantees; 22 of them are/or were also START grantees). The interviews especially highlight the potential that the START programme had for preparing researchers for an ERC grant. It is not uncommon for START grantees to receive an ERC grant (e.g. ERC consolidator grant) at later stages of their career. Two interviewees reported that their ERC application was based on results of the START project. The obligation to apply for an ERC grant at the same time also had an added value for researchers less familiar with international projects (e.g. as was the case with humanities). The control and comparison groups too, see the START programme as a good instrument of preparation for an application for an ERC grant. Hence, the programme contributed to the attractiveness of Austria as a research location, as the START grant had to be carried out within an Austrian research institution (unlike the ERC, which is "portable"; i.e. the researcher can choose any research institution in Europe). The START programme, even though internationally not as well-known as the ERC starting Grant, was seen as a valuable alternative to ERC.

5 Discussion

5.1 Using a mixed methods approach, including a quantitative impact evaluation

The evaluation of the START programme used a mixed methods approach which combined qualitative and quantitative data as well as different data collection and analysis methods. This enabled the analysis to overcome the limitations of each single method while at the same time analysing the effects of the programme on different impact levels: This is sought in most evaluations analysing the effects of research funding programmes (Böhmer et al. 2008; Melin and Danell 2006; Nedeva et al. 2012; European Commission 2014a). However, a research question is not looked at from different perspectives in all studies; rather the triangulation of data is limited to the use of different data sources for different questions.

From a methodological point of view the triangulation of data sources shows its full strengths in the analysis of the career development of the grantees. A mere survey of opinions of the funded researchers asking the question how the programme has affected and impacted a certain domain, might result (as in the case of our evaluation) in an overly enthusiastic assessment of the effects of the programme. However, by combining the

[13] This obligation was introduced as both programmes share a lot of common features. In order to not infringe upon the principle of no-double funding at the EU and national level, the double application was found as a way forward. In case both applications are successful, the START funding is not granted. However, the successful candidate can keep the title of "START grantee" and receives a supplementary funding for approx. 1 year that tops up the slightly less well funded ERC.

different sources (groups of researchers) and analysis methods (online survey and case studies) we finally had to conclude that the positive career development of the START grantees cannot be directly attributed to the programme, even though the qualitative analysis suggests a certain contribution, i.e. funded researchers state that the START programme was decisive for their career development, while unsuccessful applicants were of the opinion that their careers would have developed faster with the START grant. This conclusion is different from that of some studies that investigated the impacts of similar research funding programmes (Huber et al. 2015; Gerritsen et al. 2013; Böhmer et al. 2008). These studies concluded that the funded researchers did better than the group of comparable researchers and that the positive career development of the grantees was caused by the funding or grant. The different conclusions might be explained by the START programme or with the specificities of the Austrian science system. It is however likely, that the different method used, in the case of our study a control group approach instead of a comparison group (consisting of unsuccessful START candidates), was the reason for the different result. Our approach controlled for further variables such as discipline, scientific age and sex and randomly selected a sample of a group of similar Austrian researchers. Furthermore, a bias resulting from a too large proximity to the START programme and the FWF as funding body was reduced, as most of the researchers in the sample had never applied for the START programme and, in some cases, did not even know the programme.

The robustness of the study results also benefitted from confronting different sources of evidence. In the case of the host institutions, the views of grantees and their host institutions differed to some extend with regards to the quantity and quality of support the host institution provided and the grant holder received. One explanation for the different perceptions could be the fact that Austrian host institutions have only in recent years increased and formalised their support towards START candidates and START grantees. The different supporting activities, such as preparation courses for START applicants might have not yet been known to the actual START grantees.

5.2 Attribution versus contribution

In order to be able to make statements concerning the attribution of the programme's effects the evaluation focused on impacts on the funded researchers. This is also commonly done in other studies. The results of our study are hence similar to the analysis of similar programmes (European Commission 2014a; Melin and Danell 2006) and showed that the scientific performance, especially the citation rate and international collaboration, was better within the group of grantees than in the control groups. Next to the quantitative figures, the combination of quantitative and qualitative data collection and analysis methods proved to be very useful for explaining the impacts the programme had on the scientific performance of the grantees. While the bibliometric analysis clearly showed an increase in the scientific performance over time and a significant difference with regards to the control group, the case studies provided explanations of why the START programme produced effects e.g. the length and security of funding as an impetus to access new fields of research.

When it comes to measuring the effects of a programme on the groups consisting of indirect programme beneficiaries (here referred to as the group members or host institutions at the meso level and systemic effects at the macro level), those effects are less quantifiable. Moving up the impact chain, establishing a direct causal link between effects and the programme becomes increasingly more difficult (Nedeva et al. 2012). This is why

this study works with degrees of plausibility, rather than claiming certainty when it comes to drawing conclusions. In this sense, the evaluation is aimed at "assessing confidence" about the impact of the START programme rather than measuring the impact (Befani and Mayne 2014) and setting out the "contribution story" (Mayne 2012). With this in mind, a comparison with similar programmes outside of Austria (e.g. ERC or Emmy Noether) was not deemed useful, as START grants go mostly to Austrian researchers or researchers with some prior knowledge of the Austrian science system (e.g. through a PhD or post-doc phase in Austria). It is therefore a national programme with its national specificities which have to be taken into account in the analysis of impacts. This was done via the qualitative interviews with researchers, host institutions, group members and funders and policy stakeholders.

5.2.1 Apart from methodological constraints, fixed time and budget specifications

Prevented a stronger focus on the meso and macro levels. In this sense, the chosen evaluation design is tailor-made for the information needs of the commissioning body and takes into account the national context in which the funding programme is embedded, as recent evaluation guidelines suggest (Guthrie et al. 2013). With regards to the effects on the group members, it would have been valuable to look more closely at the future career development of the group members, especially after their time in the START group. This could for example have been done with CV analysis methods (Canibano and Bozeman 2009) in combination with a survey of former group members. Shedding more light on the effects of the START programme on the Austrian research system, especially the way basic research and research funding is perceived in the Austrian public, could have been of interest. As the START programme awards five to ten persons per year in a publicised event, the programme also has the potential to raise awareness of science in the general public. Those indirect effects on the public opinion could have e.g. been analysed through an analysis of media articles or the digital footprints of the START grantees but the framework conditions of a commissioned study did not allow a broader approach which could also investigate broader impacts.

6 Conclusion

In October 2014, the Fraunhofer ISI and KMU Forschung Austria were commissioned by the Austrian Science Fund FWF to evaluate the START programme, created in 1996. The evaluation used a mixed-methods approach with quantitative and qualitative data and analysis methods. The core methods used were a bibliometric analysis of scientific outputs of the START grantees with a control group; online surveys filled out by the START grantees, the control group and unsuccessful applicants; interviews and case studies as well as literature research and an analysis of monitoring data. A validation workshop complemented the approach.

From the point of view of a research impact assessment it has to be mentioned that this evaluation focussed on the direct beneficiaries of the programme, the START grantees, for whom the strongest effects were expected, in terms of an increase in their scientific performance and in their qualifications for leading positions in the (Austrian) research system. The second group, who is targeted by the START programme, consists of the members of the START project groups. These young researchers profit from the guidance

of top-level researchers. In the long run, they are expected to perform internationally recognised top level research themselves and widen the human resource pool of the research system in Austria. The third group targeted by the START programme is the general public in Austria. By providing funding to individual researchers the possibility of communicating and explaining research to a broader public arises, thereby increasing the acceptance of research in general and the necessity of its public funding. In a longer time perspective, the START programme is expected to contribute to strengthening the Austrian research system, which is recognised as a well-performing and attractive system by researchers and research stakeholders in- and outside Austria. However, given the time and resource constraints of a contract study, it was not possible to investigate these broader impacts by a mixed-methods approach but only to collect qualitative evidence.

The START programme shows unique features within the Austrian funding landscape; it is the only programme in Austria which provides a sort of "starting package" for a scientific career. The evaluation showed that the START programme had considerable positive effects on the scientific performance of the START grantee: they performed significantly better than the control group for nearly all bibliometric indicators measured. There is also evidence that the START programme enables testing new and unconventional research fields. Additionally, it strongly contributed to the career development of the START grantees. All START grantees remained in the science system; the majority in an Austrian institution. START project groups were also effective instruments for qualifying young researchers. Throughout its 20 years of existence, it has shaped the Austrian basic research landscape and has contributed to the strengthening of capacities and its attractiveness. Based on the numerous positive impacts for the Austrian science system, the recommendation of the evaluation team to the FWF is to continue the programme.

References

Archambault, É., VignolaGagné, E., Côté, G., Larivière, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics, 68*(3), 329–342.

Befani, B., & Mayne, J. (2014). Process tracing and contribution analysis: A combined approach to generative causal inference for impact evaluation. *IDS Bulletin, 45*(6), 17–36.

Bogner, K., & Landrock, U. (2015). *Antworttendenzen in standardisierten Umfragen*. Mannheim, GESIS—Leibniz Institut für Sozialwissenschaften (SDM Survey Guidelines). doi:10.15465/sdm-sg_016.

Böhmer, S., & Hornbostel, S. (2009). *Postdocs in Deutschland*: *Nachwuchsgruppenleiterprogramme im Vergleich*. Berlin: iFQ-Working Paper 6. http://www.forschungsinfo.de/publikationen/workingPaper.php#2009. Accessed 10 Aug 2016.

Böhmer, S., Hornbostel, S., Meuser, M. (2008). *Postdocs in Deutschland*: *Evaluation des Emmy Noether-Programms*. Berlin: iFQ-Working Paper, 3. http://www.forschungsinfo.de/publikationen/download/working_paper_3_2008.pdf. Accessed 16 Aug 2016.

Canibano, C., & Bozeman, B. (2009). Curriculum vitae method in science policy and research evaluation: The state-of-the-art. *Research Evaluation 18*(2), 86–94. http://rev.oxfordjournals.org/content/18/2/86.full.pdf#page=1&view=FitH. Accessed 20 Sept 2016.

Chi, P.-S. (2013). Do non-source items make a difference in the social sciences? In *Proceedings of ISSI 2013—the 14th international conference of the international society of scientometrics and informetrics,* Vienna, Austria, 7/15/2013.

Conchi, S., & Michels, C. (2014). Scientific mobility—An analysis of Germany, Austria, France and Great Britain. Karlsruhe. *Fraunhofer ISI Discussion Papers Innovation Systems and Policy Analysis* (41), http://www.isi.fraunhofer.de/isi-de/p/publikationen/diskpap_innosysteme_policyanalyse.php. Accessed 10 Aug 2016.

European Commission (Ed.). (2014a). *Marie Curie researchers and their long-term career development*: *A comparative study*. Written by Economisti Associati. Brussels. https://ec.europa.eu/research/

evaluations/pdf/archive/other_reports_studies_and_documents/marie_curie_researchers_and_their_long-term_career_development_-_a_comparative_study.pdf#view=fit&pagemode=none. Accessed 26 Aug 2016.

European Commission (Ed.). (2014b). *Study on assessing the contribution of the framework programmes to the development of human research capacity*. Prepared by IDEA Consult; iFG; PPMI. https://ec.europa.eu/research/evaluations/pdf/archive/other_reports_studies_and_documents/fp_hrc_study_final_report.pdf#view=fit&pagemode=none. Accessed 26 Aug 2016.

FWF (Ed.). (2007). *A contest between nations; or how far is Austrian research behind that of the world leaders? An analysis on the competivness of Austria's scientific research in the natural and social sciences*.

Gerritsen, S., Plug, E., & van der Wiel, K. (2013): Up or out? How individual research grants affect academic careers in the Netherlands. In *CPB Discussion Paper*, 249. The Hagues https://www.cpb.nl/sites/default/files/publicaties/download/cpb-discussion-paper-249-or-out-how-individual-research-grants-affect-academic-careers-netherlands.pdf. Accessed 20 Sept 2016.

Grubel, H. G. (1994). Brain drain, economics of. In T. Husen & T. Neville Postlethwaite (Eds.), *The international encyclopedia of education* I (pp. 554–561).

Guthrie, S., Wamae, W., Diepeveen, S., Wooding, Steven, & Grant, Jonathan. (2013). *Measuring research. A guide to research evaluation frameworks and tools*. San Francisco: RAND Corporation.

Hornbostel, S., Böhmer, S., Klingsporn, B., Neufeld, J., & von Ins, M. (2009). Funding of young scientist and scientific excellence. *Scientometrics, 79*(1), 171–190. doi:10.1007/s11192-009-0411-5.

Huber, N., Wegner, A., & Neufeld, J. (2015). *MERCI (Monitoring European Research Council's implementation of excellence): Evaluation report on the impact of the ERC Starting grant programme*. Berlin: iFQ-Working Paper 16. http://www.forschungsinfo.de/Publikationen/Download/working_paper_16_2015.pdf. Accessed 20 Sept 2016.

Joly, P.-B., Gaunand, A., Colinet, L., Larédo, P., Lemarié, S., & Matt, M. (2015). ASIRPA. A comprehensive theory-based approach to assessing the societal impacts of a research organization. *Research Evaluation 24*(4), 440–453. doi:10.1093/reseval/rvv015.

Krumpel, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality and Quantity: International Journal of Methodology, 47*(4), 2025–2047.

Leeuw, F., & Vaessen, J. (2009). *Impact evaluations and development: NONIE guidance on impact evaluation*. Network of Networks on Impact Evaluation (NONIE).

Lowell, B. L. (2002). *Skilled labour migration from developing countries: Annotated bibliography*. International Migration Papers, 56, International Labour Office, Geneva. http://www.ilo.org/public/english/protection/migrant/download/imp/imp56e.pdf. Accessed 20 Sept 2016.

Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation, 16*(1), 1–24.

Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation, 18*(3), 270–280. doi:10.1177/1356389012451663.

Melin, G., & Danell, R. (2006). The top eight percent: Development of approved and rejected applicants for a prestigious grant in Sweden. *Science and Public Policy, 33*(10), 702–712.

Merton, R. K. (1973). *The sociology of science*. Chicago: The University of Chicago Press.

Meyer, N., & Bührer, S. (2014). *Impact evaluation of the Erwin Schrödinger fellowships with return phase*. Final Report for the Austrian Science Fund (FWF). Karslruhe. https://ec.europa.eu/research/evaluations/pdf/archive/fp7-evidence-base/national_impact_studies/impact_evaluation_of_the_erwin_schroedinger_fellowships_with_return_phase.pdf. Accessed 26 Aug 2016.

Morton, S. (2015). Progressing research impact assessment. A 'contributions' approach. *Research Evaluation 24*(4), 405–419. doi: 10.1093/reseval/rvv016.

Nedeva, M., Braun, D., Edler, J., Frischer, D., Glanz, M., & Gläser, J. et al. (2012). *Understanding and assessing the impact and outcomes of the ERC and its funding schemes (EURECIA)*. Final Synthesis Report.

Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. *Journal of Informetrics, 1*(2), 161–169.

Penfield, T., Baker, M. J., Scoble, R., & Wykes, M. C. (2014). Assessment, evaluations, and definitions of research impact. A review. *Research Evaluation, 23*(1), 21–32. doi:10.1093/reseval/rvt021.

Preisendörfer, P., & Wolter, F. (2014). Who is telling the truth? A validation study on determinants of response behaviour in surveys. *Public Opinion Quarterly, 78*(1), 126–146.

PREST (Ed.) (2002). *Assessing the socio-economic impacts of the framework programme*. Manchester. https://ec.europa.eu/research/evaluations/pdf/archive/other_reports_studies_and_documents/assessing_the_socio_economic_impacts_of_the_framework_programme_2002.pdf. Accessed 16 Aug 2016.

Reckling, F. J., & Fischer, C. (2010). Factors influencing approval probability in FWF decision-making procedures. FWF Stand-Alone Projects Programme, 1999 to 2008. *FWF Discussion Paper*.

Roessner, D. (2000). Quantitative and qualitative methodes in the evaluation of research. *Research Evaluation, 8*(2), 125–132.

Seus, S., Heckl, E., & Bührer, S. (2016). *Evaluation of the START Programme and the Wittgenstein Award*. doi:10.5281/zenodo.50610. https://zenodo.org/record/50610/files/Eval-START-Witt_final_report.pdf. Accessed 13 Feb 2017.

Spaapen, J., & van Drooge, L. (2011). Introducing 'productive interactions' in social impact assessment. *Research Evaluation, 20*(3), 211–218. doi:10.3152/095820211X12941371876742.

van Arensbergen, Pleur (2014): *Talent Proof Selection Processes in Research Funding and Careers*. Dissertation. Den Haag: Rathenau Instituut.