A statistical evaluation of the effects of a structured post-doctoral programme

1) Alexey Bessudnov (corresponding author). Department of Sociology, Centre for Advanced Studies, National Research University Higher School of Economics. 3 Kochnovsky proezd Moscow 125319 Russia. +74991520361. abessudnov@hse.ru.

2) Igor Guardiancich. Centre for Welfare State Research, Department of Political Science, University of Southern Denmark. Campusvej 55, Odense M, DK-5230, Denmark. +4521261623. igu@sam.sdu.dk.

3) Ramon Marimon. European University Institute, UPF – Barcelona GSE, CEPR and NBER. Villa La Fonte, 20 Via delle Fontanelle San Domenico di Fiesole 50014 Italy. +39 055 4685809. ramon.marimon@eui.eu.

"This is an Accepted Manuscript of an article published by Taylor & Francis in the Studies in Higher Education on 30 April 2014, available online: **http://www.tandfonline.com/**10.1080/03075079.2014.899340*."*

Abstract

Post-doctoral programmes have recently become an important step leading from doctoral education to permanent academic careers in the social sciences. In this paper we investigate the effects of a large and structured post-doctoral programme in the social sciences on a number of academic and non-academic outcomes of fellows. We employ propensity score matching to match fellows with applicants with similar characteristics who did not receive the fellowship, and then compare the outcomes in the treatment and control groups. The programme has a statistically significant positive effect on the general life satisfaction of former fellows and their publication activity. We argue that an active and collegial research environment, with training in academic skills during post-doctoral employment, may improve the academic outcomes of post-doctoral fellows.

Keywords: educational evaluation, post-doctoral programme, academic career, social sciences, propensity score matching.

Word count: about 7700

1. Introduction

In 2006, about 200,000 doctoral degrees were awarded in the OECD countries (Auriol 2010). The average annual number of PhD degrees awarded in the USA in 2000-2006 was about 42,000 (Chiswick et al. 2010, Nerad 2004). Traditionally, doctoral education has been seen as preparation for a career in academia. To a large extent, this still continues to be true, although some doctoral graduates intentionally choose other career paths. However, as many recent PhD graduates know from their own experience, the transition from the graduate school to academic employment may be full of uncertainty and a job is not guaranteed. The crisis of postgraduate education is widely discussed both within and outside of academia (Benton 2009, Economist 2010, Youn 2005). It is often claimed that universities produce too many PhDs, while the number of academic positions is quite limited. Post-doctoral positions partially fill this gap between the supply of PhDs and the demand in the job market; for those who want to follow a research/academic career, they are an attractive alternative to temporary low-paid teaching positions. Although practical arrangements vary across countries and disciplines, post-doctoral positions are usually temporary academic positions (with the length of the contract ranging from one to five years), focused primarily on conducting research, either individually or as part of a laboratory or research group. Is the role of post-doctoral programmes to 'buy time' while searching for a job that would satisfy expectations and make proper use of the acquired credentials?

Unfortunately, there does not seem to be an answer to this first question: as Neumann and Tan (2011) put it, 'historically, the transition from degree to career is recognized as an area

neglected by research.' Nevertheless, within this area of neglect, there are several studies that look at the career development of PhD graduates in a number of countries, while our work focuses specifically on the role of post-doctoral programmes.

Nerad (2004) reviews the situation of doctoral education in the USA with data from three surveys of PhD students and graduates conducted in the USA: the Survey on Doctoral Education and Career Preparation, the 2000 National Doctoral Program Survey and the 'PhDs – Ten Years Later' survey. The latter survey found that 10 to 14 years after the completion of the PhD, about two thirds of graduates in mathematics, English and political science were employed as faculty, while for biochemistry this number was about a half and for computer science and electrical engineering about one third. This clearly reflects different career paths across academic disciplines, and wider availability of jobs in industry in more technically oriented fields. However, even for the graduates who eventually did find academic positions, the transition from graduate school to permanent employment was not easy. According to Nerad (2004), in English, political science and mathematics, it took PhD graduates up to four years after degree completion to find a stable job. Before this, many were employed in temporary and low-paid teaching positions[1].

In a later survey, 'Social Science PhDs – Five Plus Years Out (SS5)', Morrison et al. (2011) specifically looked at the careers of PhD graduates in the social sciences. Their survey sample included about 3,000 PhD graduates from US universities who completed their PhD between 1995 and 1999. The survey found that 73% of the graduates who wanted to remain in academia eventually found tenure-track positions. Many PhD graduates worked in the non-

academic sector. In both sectors, respondents reported a high level of job satisfaction. However, for many this was achieved only after a transitional period of insecurity and uncertainty following completion of the PhD. The picture is sombre in continental Europe. According to the estimates by Auriol (2011), in 2007-2009 – that is, before the Euro crisis – in Germany, Spain, Belgium and some other European countries more than 40% of graduates who received their PhD in the five years preceding the survey had jobs with temporary contracts.

In this context of the academic labour market it is particularly relevant to study the role of post-doctoral programmes in the early careers of PhD holders. In many disciplines, especially in the natural sciences, being in a post-doctoral position has now become a necessary step in an academic career. In 1999, there were about 40,000 post-docs in the USA (Mervis 1999). The number of post-doctoral positions differs across academic disciplines. As Nerad and Cerny (1999) reported on the basis of the 'PhD: Ten Years Later' study, post-docs were a norm in biochemistry, while in mathematics there were fewer temporary positions available. In the social sciences, post-docs are still quite rare, as only 9% of PhD graduates in the social sciences in US universities held post-doctoral positions (Morrison et al. 2011), although these are rapidly expanding in response to the supply-demand mismatch of new PhDs in the social sciences.

Post-doctoral appointments are often supposed to be a temporary stage in one's academic career, a relatively short period of time when researchers have an opportunity to strengthen their publication record and then move on to a permanent academic position. In reality, post-doctoral employment may also be a period of considerable stress. As Akerlind (2005, 2009) noted in the study of post-doctoral fellows in Australia, many post-docs are

sceptical about finding a permanent position. For some, especially in medical and agricultural research, post-doctoral positions are not a step towards a more stable academic position, but rather a career in themselves. Many post-docs reported a sense of isolation. The frequent experience of isolation of post-docs was also reported in the study by Nerad and Cerny (1999). In a survey of Canadian post-docs Helbing et al. (1998) found that the most important stress factors for the respondents were concerns about the availability of future employment and the pressing need to publish the results of their research.

Most post-doctoral programmes, especially in the social sciences, seem to give an affirmative question to our first question: they basically provide time for independent research (or research within an existing project) and do not include any training component useful to fellows when searching for a permanent academic position (such as, for example, training in research, teaching and presentation skills). This brings us to our second question: is there any added value in having a structured post-doctoral programme aimed at improving these skills?

In educational studies, there is a large literature that addresses the development of academic skills among recent doctoral graduates. It is generally agreed that in most cases doctoral education does not provide all the skills necessary for a successful academic career. Although, according to Gardner's model (Gardner 2008), successful academic socialization in the graduate school already implies the formation of the skills required for independent research, many young academics find further training and mentoring in the early stages of their careers useful. Furthermore, as already noticed, post-doctoral programmes seldom are an example of successful academic socialization.

On the basis of a qualitative study, conducted in two research oriented UK business schools, Adcroft and Taylor (2013) suggest four crucial spheres for career support of young academics: managing expectations, career management, mentoring and professional development. Given that there is a lot of tacit and field specific knowledge about what is required for a successful academic career, managing expectations of newcomers to academia is important, and mentoring by senior colleagues, both formal and informal, becomes almost invaluable. This is often supplemented by formal professional development courses, frequently supplied at the institutional level. In terms of career management, many young academics find it hard to combine teaching and research.

A large literature looks at the development of teaching skills of young faculty, conducted in the form of various types of interventions (workshops, seminars, one-to-one consultations, etc.). These interventions proved to be quite effective (for recent systematic reviews see Amundsen and Wilson 2012; Stes et al. 2010). In this context, Knight et al. (2006) stress the importance of non-formal learning; Hubball and Poole (2003), on the other hand, emphasize structured formal programmes of study.

Traditionally, the field of academic development mainly focused on the development of teaching skills and less so on on research skills (Akerlind 2008). However, given the emphasis that modern universities make on research conducted by the faculty and their publication activity, developing research-related skills also becomes crucial for an academic career. Writing skills are an important part of wider research skills. Murray and Cunningham (2011) describe an intervention (writer's retreat) aimed to improve publication activity of young faculty in an

Australian university. The important features of this intervention was its structured character (participants were given a schedule), work in groups (participants were writing in the same room) and simultaneous focus on developing writing projects and participation in research assessment. Aitchison (2009) and Lee and Boud (2003) describe the experience of writing groups more generally. McGrail et al. (2006) provide a systematic review of interventions aimed to improve the publication rates of the faculty and conclude that all forms of intervention (writing courses, writing groups and writing coaches) led to an increase of publication rates. Reid and Petocz (2003) stress the importance of the simultaneous development of research and teaching skills.

A conclusion from this literature that is important for our study is that many successful interventions in academic development have a structured nature and focus on several skills required for an academic career at the same time. Adcroft and Taylor (2013) emphasize that in order to succeed in career development of young academics, different types of support should be provided together and simultaneously.

A structured approach to academic development, focused on teaching, research and other types of skills of participants at the same time, is a key feature of the Max Weber Programme (MWP) at the European University Institute (EUI) that we study in this paper. We are interested in the long-term effects such a programme has on a number of academic and non-academic outcomes of fellows. In particular, our contribution in this paper is the following: we assess the effectiveness of a structured post-doctoral programme statistically, following the experimental and quasi-experimental tradition that in the last decade has gained firm ground in

the programme evaluation literature (Imbens and Wooldridge 2009). The idea is that we compare a number of academic and non-academic outcomes in the 'treatment' group of MWP fellows and the control group of applicants who applied for the fellowship, but who were not selected. To provide unbiased estimates of the differences in various academic outcomes between these two groups, we employ propensity score matching to match them according to several pre-treatment characteristics (i.e., the characteristics known before the selection for the programme was made).

A similar design was previously used to evaluate the effectiveness of the National Academy of Education/Spencer post-doctoral fellowship programme in the USA (Hedges et al. 2011). This, to our knowledge, is the only other study that has statistically studied a post-doctoral programme. Instead of propensity score matching, Hedges et al. – counting with a more continuous assessment variable – applied the regression discontinuity design, although the underlying logic of selection on observables was the same as in our study. They found a positive effect of the Spencer fellowship on a number of academic outcomes, such as the number of publications, influence on the work of others and the number of grants received. Our results show that the MWP fellows – inserted in a collective and structured programme – have higher general life satisfaction and a higher score on the publications index, as compared to the scholars in the control group. The effect on satisfaction with teaching is less robust and statistically significant only at the 90% level. The results indicate that training of fellows that aims to improve their teaching, writing, publishing and other academic skills has a positive effect on their further academic careers.

The paper proceeds as follows. In section 2 we describe the MWP in more detail,

with a focus on the training it provides. Section 3 describes the data collection procedure and section 4 explicates the statistical methods we use. Section 5 provides the results of the analysis, first showing the effects of the MWP and then estimating potential effects of survey non-response. In sections 6 and 7 we discuss the results substantively and conclude.

2. The Max Weber Post-doctoral Programme

The Max Weber post-doctoral programme of the European University Institute, funded by the European Commission (DG Education and Culture), is the largest post-doctoral programme in the social sciences in Europe and, possibly, the only one that is structured. It began in the academic year 2006-2007, and since then about forty fellows are admitted to the programme every year, usually for a one-year fellowship. The fellows are associated with one of the four EUI departments: Economics, Social and Political Science, History, and Law. Admission to the programme is highly competitive, with an average acceptance rate of 5.4% for the early years of the study and under 4% in more recent years. Max Weber fellows are selected by the faculty of the different EUI departments, together with the MWP Director, on the basis of: i) research accomplishments and potential (publications, PhD thesis, projects, recommendation letters, etc.); ii) academic career interests (priority given to those who plan to pursue an academic career), and iii) availability of EUI faculty to provide mentorship (close match of research or teaching interests is not required).

The MWP has developed a new model for a structured post-doctoral programme, which leaves time for fellows to develop their own research agenda – actively participating in the research activities of their respective departments – while providing them with a range of

activities to enhance their academic skills in a stimulating multicultural and multidisciplinary environment. Namely, the MWP offers a range of activities for its fellows, which can be grouped under Multidisciplinary Research Activities and Academic Practice Activities. The former are designed to enhance the communication and understanding of current research across disciplines and, since most of them (workshops and conferences) are organized by the fellows themselves this also helps them to broaden their scholarly skills. The latter focus more specifically on improving academic skills and cover a range of activities, which in turn can be grouped under: presentation and communication; academic writing (publishing and grant applications); teaching; academic job market, and others (academic ethics, international comparison of academia and academic careers, etc.). These take the form of individual tutoring by academic staff, group discussions among fellows, workshops and conferences. These activities are often the basis of daily informal discussions among the fellows, who share common spaces, meals, etc.

While all fellows participate in some of these activities (e.g. presenting their work to all the other fellows, writing a grant proposal and at least a working paper, and designing a new course), fellows participate in others according to their needs and interests, but also participate in giving feedback to other fellows. For example, most fellows actively participate in 'academic job market' activities (e.g. to improve their performance in interviews) but, through the year, also actively participate in the preparation of job-market seminars of other fellows. Similarly, fellows tend to have different experience and needs regarding teaching and, although the programme does not require teaching, fellows can participate in a number of workshops, given by educational experts collaborating with the programme, give seminars for PhD researchers at the EUI, teach courses at BA and MA levels at the many American and Italian universities in

Florence with which the MWP has collaborative agreements, as well as embark on one-week teaching experiences, with feedback, at either the London School of Economics, the Humboldt University in Berlin, or the Universidad Pompeu Fabra in Barcelona.

In sum, the Max Weber Programme is a structured post-doctoral programme that covers most aspects related to successful advancement in an academic career[2]. An obvious question arises: does it make a difference for those who participate in it?

3. Data

The data for this evaluation study come from two sources. First, we collected data from the applications sent by the applicants to the MWP. These data include gender, age, nationality, the name of the university that awarded the PhD degree, the assessment of the 'quality' of the candidate summarized by the director of the MWP from the joint selection process with the faculty of the EUI, the year of the application, the department, and a dummy variable indicating whether the candidate was selected for the programme. These data were not available for the 2006-2007 applications. For 2007-2008 the missing data on the 'quality' of the candidates was assessed by the director of the programme *post hoc* (but before conducting the survey and before running any statistical analysis). Overall, we collected application data for three waves of applicants (2007-2008, 2008-2009 and 2009-2010).

The second data source is a survey conducted among the former Max Weber fellows and the applicants that were not selected for the programme. It was neither practically possible nor methodologically meaningful to conduct a survey among all unsuccessful applicants, including the candidates with very little chance of being selected for the MWP. To choose the

applicants for the survey, we employed a statistical procedure of matching to select a group of unsuccessful applicants who most closely resembled the group of fellows.

The survey was conducted from June to September 2011 via the Internet (on Limeservice.com) and included questions about the current job of the respondents, the type of contract they hold, funding opportunities, satisfaction with different aspects of their academic careers and life in general, their publication record, etc. The invitations to take part in the survey were sent by email, with up to five reminders. The email addresses of the fellows and the applicants were found on the Internet. Overall, 117 invitations were sent to fellows and 261 to non-fellows. The response rate for the group of fellows was 68% and for the control group of other applicants 35%. Table 1 reports the descriptive statistics for both data sources, including the response rates for the survey.

[Table 1 near here]

4. Methods

A simple comparison of the group of fellows and the group of other applicants would give a biased estimate of the effect of the MWP on the academic outcomes of the fellows. It is obvious that the fellows were selected for the programme on the basis of their academic abilities and we would expect them to have better academic results even without the MWP. However, it is possible to find some applicants in the group of non-fellows who closely resemble the fellows in their observed characteristics, but, for particular reasons (such as, for example, the absence of a suitable mentor), were not selected for the programme. We use these applicants as a control group for the fellows.

To select the control group, we applied propensity score matching (Rosenbaum and Rubin 1983, Rosenbaum and Rubin 1984, Dehejia and Wahba 2002). At the first stage, we employed matching to select a group of applicants for the survey. We used the following 'pre-treatment' variables (available in the applications) to match the fellows and non-fellows: gender, department, nationality, the rank of the university that awarded a Ph.D. degree according to the QS Social Science University ranking (the universities were divided into five groups according to their rank), the country of the university, and the quality of the applicant as assessed by the director of the MWP. We estimated the propensity scores for each year separately. A propensity score is a probability of being selected for the programme, given the 'pre-treatment' variables listed above. The applicants who were not selected for the programme and whose propensity scores were high enough to be compared with the scores of the fellows were chosen for the survey.

Note that the characteristics of the applicants used to estimate the propensity score was the only information available to the MWP to select the successful candidates. No personal interviews were conducted, and the selection procedure was based entirely on a CV, cover letter, and references. Thus, there are no unobserved factors that remain unaccounted for in the model and that could affect the selection procedure.

After the survey was conducted, we estimated the propensity score again, this time using only those fellows and members of the control group who completed the survey. The variables used for matching were the same as before, although this time we estimated one model for all years and added the year as a matching variable. We excluded all the applicants who were

selected for the programme, but declined the offer (fifteen individuals among those who completed the survey). These were mostly economists and too few for a statistically meaningful analysis.

After we estimated the propensity score, we employed several techniques to find the average treatment effect of the MWP. To compare the 'experimental' and control groups, we first used simple linear regression, without taking into account the propensity score. Second, we used regression that controls for the propensity score. Third, we applied the stratification matching estimator. All these techniques were applied only to the observations in the region of common support as estimated by the propensity score matching (i.e. in the part of the distribution where the propensity scores in the group of fellows and non-fellows overlapped). The statistical analysis was conducted in Stata (Becker and Ichino 2002).

5. Results

5.1. Effects of the MWP

Table 2 shows descriptive statistics for 'pre-treatment' variables (i.e., the variables available before the survey) in the treatment and control groups. The two groups are fairly well matched, although it should be noted that there are more applicants whose quality was assessed as 'best' in the group of fellows, and more applicants who were just 'adequate' in the group of matched applicants. The control group also has fewer lawyers than the group of fellows, and more economists. Applicants who obtained a PhD in French universities are more numerous in the group of fellows, while there are more applicants with degrees from the USA in the control group.

[Table 2 near here]

Table 3 demonstrates the distribution of the propensity score in the treatment and control groups (in the matched sample in the common support region). Quite naturally, fellows have a higher mean propensity score. However, note that in further statistical analysis we compare the fellows and non-fellows only in the part of the distribution where propensity scores in two groups overlap. Also, we use stratification matching to split the distribution into five parts where the mean propensity score in two groups is comparable.

[Table 3 near here]

Table 4 shows the effects of being a Max Weber fellow for the matched sample, only for the variables that are statistically significant at the 95% and 90% levels. These are three variables: life satisfaction (measured on a 10-point scale), the publications index, and satisfaction with teaching (only significant at the 90% level). The full set of estimates is available in the online Appendix.

[Table 4 near here]

If we take the estimates from the stratification matching model, being a Max Weber fellow increases general life satisfaction by 44% of the standard deviation (95% CI: 8-79%).

To estimate the effect of the MWP on the quality and quantity of academic publications, we constructed a publications index based on the three best publications as reported by respondents in the survey. The publications index closely follows the methodology adopted

by the European Association of Development Research and Training Institutes (EADI). We assigned a number of points to different types of publications: journal articles, edited volumes, monographs and book chapters. Each publication falls into one of five categories (A to E). For academic journals, categories A and B are reserved for journals included in the Science Citation Index and Social Science Citation Index of the ISI Web of Knowledge (2010). All journals above the cut-off point of 33% in each sub-discipline are ranked A; those below are ranked B. Other refereed journals that are not (yet) ranked by the ISI get a C; all non-refereed journals for mainly academic readers get D, all others E. As there is no ISI ranking system for book publishers, we have used a rating system based on the visibility of a scientific publisher on Google Scholar. Publishers who score more than 500,000 hits get an A, from 50,000 to 500,000 hits – B, from 1,000 to 50,000 hits – C, from 100 to 1,000 hits – D, and less than 100 hits and unclear – E. Finally, each category is subdivided into two: whether the publication is single- or co-authored (the former yields more points). The total score is the sum of obtained points.

The effect of the MWP on the publications index is estimated to be 59% of the standard deviation (95% CI: 11-107%). However, a word of caution must be added. First, the effect is not statistically significant in one of the model specifications (regression controlling for the propensity score). Second, it is based on a limited sample (n=97) as some participants of the survey did not answer the question about their best publications.[3]

The third effect of the MWP that is, however, only statistically significant at the 90% level and only in one out of three model specifications, is satisfaction with teaching. The size of the effect in the stratification matching specification is 37% of the standard deviation (95% CI: -11-86%). We do not consider this effect to be robust.

We attempted to estimate the separate effects of the MWP for different cohorts, departments, genders, and fellows of different pre-treatment academic 'quality'. All the interaction effects between these variables and being a fellow were statistically not significant. While it is very likely that the programme affects fellows with different characteristics in a different way, our sample size is too small for a reliable estimation of these effects.

5.2. Determinants of non-response

A limitation of this study is that we do not account for the possibility of a differential non-response to the survey that may bias the results. The response rate in the group of fellows was about twice higher than in the group of non-fellows (see Table 1). If the probability of responding to the survey is associated with any characteristics that can affect academic productivity and other outcomes, then our estimates of the effect of the MWP may be biased. To test this, we looked at the association between the probability of completing the survey and several pre-treatment characteristics (such as the academic 'quality' of the applicants, gender, department, year of application and nationality), separately for fellows and other applicants. We also regressed probability of non-response on the estimated propensity score. The results are presented in Table 5.

[Table 5 near here]

For the group of fellows, there are no statistically significant predictors of non-response. There is no evidence that the 'quality' of applicants, gender, year of application, department and nationality affect the decision to take part in the survey. The model with the propensity score shows that it is not statistically significant either.

For the group of non-fellows, the 'quality' of the applicants is not a statistically significant predictor of non-response either. However, in this group the applicants from the departments of History and Social and Political Science were less likely to complete the survey compared to economists. Applicants from the USA were more likely to complete the survey than applicants of other nationalities. As follows from Table 2, this leads to some oversampling of the economists and applicants from the USA in the matched group of non-fellows. However, the oversampling is small and it is unlikely to seriously bias the results. Of course, it is possible that there are some other unobserved factors that simultaneously affect non-response and the outcomes and bias our estimates. This, however, cannot be tested empirically.

6. Discussion

How do we interpret these results? First, it must be noted that we estimate the effect of the MWP vs. the average effect of other academic positions in our sample. People in the control group who did not get the MWP fellowship were selected for other academic positions. The literature on causal inference emphasizes that the effects of treatment can only be estimated relative to the effects of other treatments that were received by units in the control group (Holland 1986).

The most robust of our findings is the higher general life satisfaction of the MWP fellows. It is tempting to explain this effect simply by the fact that fellows spent a year in Florence, Italy, almost universally acknowledged as a location with a good climate and a culturally stimulating environment. We do not accept this explanation as very convincing. It is unlikely that a year spent in Florence would have a long-lasting effect on general life

satisfaction. People tend to adapt to current circumstances in their lives, and good events and experiences often have only a short-term effect on happiness (an effect known as hedonic adaptation in psychology (Diener et al. 2011)).

Another possible explanation is that the effects are statistical artefacts. Indeed, given that we compare twenty-seven outcomes in two groups we can reasonably expect the difference in one or two outcomes to be statistically significant at the 95% level simply by chance. We found two effects that are significant at the 95% level and one that is significant at the 90% level. However, all these effects are in the expected direction (i.e., positive for the MWP fellows), and this is unlikely to happen just by chance.

It is also possible that the higher reported life satisfaction of the fellows is simply a short-term effect of the administration of the survey on the current mood of respondents. Indeed, by sending an invitation to take part in the survey to former fellows we remind them of the prestigious fellowship they won and of the year they spent in Florence. On the contrary, applicants in the control group were reminded of the not so pleasant experience of not getting a fellowship. Psychological research demonstrated that these contextual effects can affect reported life satisfaction (Kahneman and Krueger 2006). Although we cannot rule out this possibility, it does not explain the MWP effect on the publications index.

Our explanation is the long-term effect of the MWP on the academic adjustment of the fellows. At the EUI the fellows were given training in many aspects of academic life – including teaching, academic writing, publication strategies, time management, organization of workshops, etc. – and were made familiar with other academic institutions, publishers, etc.

Fellows from different disciplines, universities and countries were given opportunities to communicate and share their academic experiences, which provided them with support in their job-search, and a better perspective on the challenges faced in an academic career. We argue that this had a long-term effect on the psychological adjustment of the fellows to academic life, helped them to deal with the stress of the early stages of academic careers, and affected their general life satisfaction. It is also instructive that we identified the effect of the MWP on the publication record and on teaching satisfaction (although this effect is not robust), two fields where fellows were given especially intense training.

These findings are in line with previous research on the development of academic skills of early-career academics reviewed in the Introduction. Previous studies, mainly based on qualitative methodology, argued that it is the structured and systematic approach to academic development that makes it most effective (Adcroft and Taylor 2013, Hubball and Poole 2003). Although in the present design we were unable to test quantitatively the impact of the different types of activities offered at the MWP on the academic outcomes of fellows, we believe that it is their combination, reinforced by the collective nature of the MWP experience, that serves as a mechanism for the positive effects we established statistically.

It should also be noted that the fact that we identified these three effects – the most significant one being life satisfaction – discriminates against the simple explanation that the MWP makes a difference only as a selection (signalling) mechanism. While, if this were the case, the MWP would still have value for those participating in the programme, our main interest is whether it makes a difference as a structured post-doctoral programme. However, neither life

satisfaction nor teaching satisfaction are part (and could not be part) of the selection criteria.

This study has several limitations. First, we can only estimate the short-term effects of the programme. In fact, the oldest cohort of the fellows only finished the MWP three years before the survey, and the youngest only one year before the survey. It is quite likely that the MWP has long-term effects on the performance of the fellows that we cannot capture in this evaluation.

Second, our sample size is not very large, and this leads to wide confidence intervals in our estimates. In fact, in this study we are able to reliably identify only the strongest effects of the MWP. It is possible that some of the statistically insignificant effects would have been significant if the sample size was larger.

Third, survey non-response can bias the results. As shown in section 4.2, known characteristics of applicants did not affect the probability of response for fellows and only weakly affected the probability of response for non-fellows. However, it is quite likely that applicants who left academia did not complete the survey. If there is a substantial number of applicants who left academia in the control group this would make our estimates of the effect of the MWP conservative.

In spite of these limitations, follow-up research has provided additional support to the findings reported here. In particular, Cirillo, Marimon and O'Malley (2013) have expanded the sample analysed here (including the 2010–2011 cohort and getting data from 167 former Max Weber Fellows, 248 unsuccessful 'highly ranked' candidates and 66 successful candidates who

declined the fellowship), analysing their careers based only on existing on-line information (Google-searching candidates and their institutions), which covered 98% of the former Max Weber fellows and 91% of non-fellows. They find that former Max Weber fellows – in relation to non-fellows – have a higher probability of career progression in academia, as well as of securing and maintaining an academic job; more specifically, they are more able to gain tenure-track positions that immediately follow the fellowship, which is a significant difference from the results of non-fellows on the job market.

7. Conclusion

With differences across disciplines and nationalities, doctoral studies in the social sciences have evolved from individual research, under the supervision of an advisor, to more structured PhD programmes where courses, seminars and the interaction among PhD researchers are a very important complement to, and stimulus for, individual research. In the natural and human sciences the post-doctoral stage is almost a pre-requisite for more permanent positions in academia or research centres, and the lab provides a social training aspect bridging the PhD with more permanent positions. While the post-doctoral stage is becoming increasingly common in the social sciences and humanities, most post-doctoral fellowships resemble the old PhD model of individual research and, rather than a bridge, act as a waiting room for more permanent offers (Nerad and Cerny 1999; Akerlind 2005).

With increasing demand for post-doctoral fellowships in the social sciences and humanities and stringent financial support, there is a further need to enhance the efficiency of

post-doctoral programmes. This study has addressed the question: is there any added value in having a structured post-doctoral programme aimed at improving academic skills? The Max Weber Programme is a pioneer structured programme in the SSH that, while leaving ample time to fellows to develop their own research agendas, provides academic training in a collegial multicultural and multidisciplinary environment, a sharp contrast with more standard 'individual fellowships'. Therefore, it is a good, and probably unique, candidate with which to address the above question[4]. It is not common to subject a new educational programme to statistical scrutiny in order to assess its effectiveness, as if it was a clinical trial. But this is precisely what we have attempted in this work. Even if, in contrast with clinical trials, we could not perform a purely random natural experiment – which would have required random selection in a very selective programme – we think that by using propensity score matching techniques we have come as close to it as possible with the available data.

The data come from three cohorts of Max Weber fellows and, as a control group, the candidates who would most likely have been accepted had the programme been a little more than twice its (already large) size. We have detected a statistically significant positive effect of the programme on general life satisfaction – where life for most of our subjects means academic life – and publication record. Even if most fellows participated in the programme only for one academic year, our results provide an illuminating qualified affirmative answer to the question: does it make a difference? It is a statistically significant yes (on life satisfaction and publication record). It is 'informative' since we think that the best explanation of these results lies in the, already mentioned, design of the Max Weber Programme. It is its structured nature, based on communication and collective activities of fellows, that improves their academic outcomes. The

further development of this programme – and, possibly, other structured post-doctoral programmes that may emerge – will provide more data to replicate (or contradict) and extend these results.

Endnotes

1. Other studies of the career trajectories of doctoral recipients are: Park (2007) and Haynes et al. (2009), both finding that only about a third of UK PhD graduates pursue academic careers, and Neumann and Tan (2011) who, using the Australian Graduate Destination survey, find that 44% of Australian PhD graduates were employed in higher education, and only half of these 44% (i.e., about a quarter of all PhD graduates) held academic positions.

2. For more detailed information, see:

http://www.eui.eu/ProgrammesAndFellowships/MaxWeberProgramme/Index.aspx .

3. Another possible approach to measuring the publication activity of the fellows and non-fellows would be to look at their complete list of publications on Google Scholar. However, there would be a problem of homonymy and, furthermore, most survey participants are at the beginning of their careers and most of their work may not have been indexed in Google Scholar.

4. A measure of the efficiency of a programme should also take into account its cost. During the years covered in this study the total yearly cost of the programme did not exceed 2.3 million euros, corresponding to the specific funding the EUI received from the European Commission, DG Education and Culture.

Acknowledgements

We would like to thank Donald Rubin for his encouragement to pursue this inquiry and Andrea Ichino, Prashant Loyalka, Fabrizia Mealli and two anonymous reviewers for their helpful comments.

Conflict of interests

Alexey Bessudnov was a Max Weber Fellow (2010–2011), Igor Guardiancich was the Max Weber Programme Academic Assistant for the Academic Careers Observatory (2010-2011) and Ramon Marimon is the Director of the Max Weber Programme (2006-2013).

References

Adcroft, A. and Taylor, D. 2013. Support for new career academics: an integrated model for research intensive university business and management schools. *Studies in Higher Education* 38(6): 827-840.

Aitchison, C. 2009. Writing groups for doctoral education. *Studies in Higher Education* 34 (8): 906-916.

Akerlind, G.S. 2005. Postdoctoral researchers: roles, functions and career prospects. *Higher Education Research and Development* 24: 21-40.

Akerlind, G.S. 2008. Growing and developing as a university researcher. *Higher Education* 55: 241-254.

Akerlind, G.S. 2009. Postdoctoral research positions as preparation for an academic career. *International Journal for Researcher Development* 1(1): 84-96.

Amundsen, C. and Wilson, M. 2012. Are we asking the right questions? A conceptual review of the educational development literature in higher education. *Review of Educational Research* 82: 90-126.

Auriol, L. 2010. Careers of doctoral holders: Employment and mobility patterns. *STI working paper* 2010/4 .

Becker, S.O. and Ichino, A. 2002. Estimation of average treatment effects based on propensity scores. *The Stata Journal* 2: 358–377.

Benton, T.H. 2009. Graduate school in the humanities: Just don't go. *The Chronicle of Higher Education*. January 30.

Chiswick, B., Larsen, N. and Pieper, P. 2010. The production of PhDs in the United States and Canada. *IZA working paper* 5367.

Cirillo, C., Marimon R. and A. O'Malley. 2013. Impact evaluation of Max Weber Programme in the Academic Job Market. MWP Academic Careers Observatory, European University Institute.

Dehejia, R.H. and Wahba, S. 2002. Propensity score matching methods for nonexperimental causal studies. *Review of Economics and Statistics* 84: 151–161.

Diener, E., Oishi, S. and Lucas, R.E. 2011. Subjective well-being: The science of happiness and life satisfaction. In *The Oxford Handbook of Positive Psychology,* 187–194. Oxford: Oxford University Press.

Economist. 2010. The disposable academic. *The Economist,* December 16.

Gardner, S.K. 2008. 'What's too much and what's too little?': The process of becoming an independent researcher in doctoral education. *Journal of Higher Education* 79: 326-351.

Haynes, K., Metcalfe, J. and Videler, T. et al. 2009. *What do researchers do? First*

destinations of doctoral graduates by subject. London: Careers Research and Advisory Centre.

Hedges, L., Hanis, J. and Asch, E. 2011. Statistical evaluations of the Spencer fellowships programs. In *Learning to Work Better. The Spencer Foundation's In-Depth Review of Its Signature Fellowship Programs*, 9-26. Chicago: The Spencer Foundation.

Helbing, C.C., Verhoef, M.J. and Wellington, C.L. 1998. Finding identity and voice: a national survey of Canadian postdoctoral fellows. *Research Evaluation* 7(1): 53-60.

Holland, P.W. 1986. Statistics and causal inference. *Journal of the American Statistical Association* 81: 945–960.

Hubball, H. and Poole, G. 2003. A learning-centred faculty certificate programme on university teaching. *International Journal for Academic Development* 8 (1/2): 11-24.

Imbens, G.W. and Wooldridge, J.M. 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47(1): 5-86.

Kahneman, D. and Krueger, A.B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives* 20: 3–24.

Knight, P., Tait, J. and Yorke, M. 2006. The professional learning of teachers in higher education. *Studies in Higher Education* 31 (3): 319-339.

Lee, A. and Boud, D. 2003. Writing groups, change and academic identity: Research development as local practice. *Studies in Higher Education* 28 (2): 187-200.

McGrail, M.R., Rickard, C.M. and Jones, R. 2006. Publish or perish: A systematic review of interventions to increase academic publication rates. *Higher Education Research and Development* 25: 19-35.

Mervis, J. 1999. The world of postdocs. *Science* 285(3): 1513.

Morrison, E., Rudd, E. and Nerad, M. 2011. Early careers of recent U.S. Social Science PhDs. *Learning and Teaching* 4(2): 6-29.

Murray, R. and Cunningham, E. 2011. Managing researcher development: 'drastic transition'? *Studies in Higher Education* 36 (7): 831-845.

Nerad, M. 2004. The PhD in the US: Criticism, facts, and remedies. *Higher Education Policy* 17: 183-199.

Nerad, M. and Cerny, J. 1999. Postdoctoral patterns, career advancement, and problems. *Science* 285(3): 1533-1535.

Neumann, R. and Tan, K.K. 2011. From PhD to initial employment: the doctorate in a knowledge economy. *Studies in Higher Education* 36: 601-614.

Park, C. 2007. *Redefining the Doctorate*. York: The Higher Education Academy.

Reid, A. and Petocz, P. 2003. Enhancing academic work through the synergy between teaching and research. *International Journal for Academic Development* 8: 105-117.

Rosenbaum, P.R. and Rubin, D.B. 1983. The central role of the propensity score in

observational studies for causal effects. *Biometrika* 70: 41–55.

Rosenbaum, P.R. and Rubin, D.B. 1984. Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association* 79: 516–524.

Stes, A., Min-Leliveld, M., Gijbels, D. and Van Petegem, P. 2010. The impact of instructional development in higher education: The state-of-the-art of the research. *Educational Research Review* 5: 25-49.

Youn, T.I.K. 2005. The academic job market is bad for all of us. *Academe* 91: 27–30.

Tables

Table 1. Descriptive statistics for the MWP applications and the survey

year	total n applicants	n new fellows	n invitations for other applicants	n completed for other applicants	n completed for fellows	response rate for other applicants (%)	response rate for fellows (%)
2007	446	34	41	16	22	39	65
2008	784	41	133	46	29	35	72
2009	928	42	87	29	28	33	67
Total	2158	117	261	91	79	35	68

Note. Numbers for other applicants are given excluding successful applicants who declined the offer. The response rate for fellows includes incomplete questionnaires. For one fellow in 2008 we could not find a valid email address and the invitation to take part in the survey was not sent.

Table 2. Descriptive statistics for the analytic sample in the common support region

variable	fellows	matched applicants	Total
Gender			
Men	42	42	84
Women	37	34	71
Year			
2007	22	16	38
2008	29	31	60
2009	28	29	57
Department			
Economics	18	25	43
History	21	18	39
Law	13	7	20
SPS	27	26	53
University rank			
1-25	26	28	35
26-50	9	5	14
51-150	23	17	40
151-300	9	7	16
>300 or absent	12	19	31
Country of university			
USA	20	27	47
UK	13	14	27
France	11	5	16
Italy	3	5	8
Other	32	25	57
Nationality			
USA	6	10	16
Italy	12	11	23
Germany	11	7	18
Other Western Europe	19	17	36
Eastern Europe and the Balkans	15	14	29
Other	16	17	33
‘Quality’ of the applicants			
best	38	21	59

good	18	20	38
adequate	23	35	58
n	79	76	155

Table 3. The distribution of the propensity score in the treatment and control groups

propensity score	fellows	matched applicants	Total
0.06 – 0.2	4	17	21
0.2 – 0.4	12	22	34
0.4 – 0.6	14	28	42
0.6 – 0.8	27	6	33
0.8 - 1	22	3	25
total	79	76	155

Note. The propensity scores are given for the common support region [0.06-1].

Table 4. The effects of the MWP

variable	mean	standard deviation	regression (no controls)	regression (controlling for propensity score)	stratification matching	n
publications index	10.6	5.4	**2.3 (1.1)**	1.8 (1.2)	**3.2 (1.3)**	97
life satisfaction	7.7	1.3	**0.45 (0.21)**	**0.5 (0.24)**	**0.57 (0.23)**	146
satisfaction with teaching	7	2.0	0.16 (0.33)	0.49 (0.37)	*0.84 (0.46)*	146

Note. Standard errors are in parentheses. The effects that are statistically significant at the 95% level are marked in bold, at the 90% level – in italics.

Table 5. Predictors of non-response

variable	fellows	other applicants		
	(1)	(2)	(1)	(2)
constant	**0.72 (0.24)**	**0.74 (0.10)**	**0.89 (0.15)**	**0.4 (0.05)**
propensity score		-0.06 (0.18)		-0.03 (0.16)
'Quality' (ref. best)				
good	-0.07 (0.11)		-0.02 (0.09)	
adequate	0.02 (0.12)		-0.04 (0.08)	
Male	0.02 (0.09)		-0.07 (0.06)	
Year (ref.2007)				
2008	0.11 (0.12)		-0.05 (0.09)	
2009	0.02 (0.11)		-0.04 (0.09)	
Department (ref.Economics)				
History	0.01 (0.14)		**-0.2 (0.08)**	
Law	-0.18 (0.14)		-0.14 (0.11)	
SPS	0.08 (0.13)		**-0.19 (0.07)**	
Nationality (ref.USA)				
Italy	0.02 (0.22)		**-0.43 (0.14)**	
Germany	0.05 (0.23)		-0.26 (0.15)	
Other Western Europe	-0.07 (0.21)		**-0.29 (0.13)**	
Eastern Europe and the Balkans	-0.02 (0.22)		**-0.33 (0.13)**	
Other	-0.18 (0.21)		**-0.33 (0.13)**	
n	116	115	284	270

Notes. Linear probability models. Standard errors are in parentheses. The effects that are statistically significant at the 95% level are marked in bold.