Human and social capital accumulation within Research Infrastructures: the case of CERN

Authors: Gelsomina Catalano[1] , Francesco Giffoni[1,2] ,Valentina Morretta[2]*

Affiliations:

[1] CSIL – Centre for Industrial Studies, Corso Monforte 15, 20122, Milan, Italy

[2] Department of Economics, Management, and Quantitative Methods, University of Milan, Via Conservatorio 7, 20122, Milan, Italy.

*Corresponding author: valentina.morretta@unimi.it

Abstract[1]

The contribution to human and social capital accumulation is one of the most important socio-economic benefits of public investment in Research Infrastructures. Sure enough, these large scientific enterprises are exceptional incubators of human and social capital, especially for early-career researchers who have the opportunity to gain new skills and expand their network of contacts in highly prestigious and challenging workplaces. This paper explores the contribution of spending a period of study and/or work at the Large Hadron Collider of CERN to the expected future lifelong salary of early-career researchers. Previous studies are here extended by using three sources of data: primary data collected through a survey to CERN Alumni, a survey to team leaders who supervised early-career researchers at CERN, and secondary data salary information. Findings show that an experience-based learning process at CERN is instrumental in developing skills that are needed by the economy and reveal an expected salary premium between 5% and 11%. Such human capital effect seems more important than a pure networking and reputational effect.

[1] **Acknowledgments:** We are grateful to M. Florio (University of Milan) and J. Gutleber (secretary of the FCC study, member of the host-state concertation structure, CERN Accelerator and Technology Sector Directorate Office) for their suggestions and useful feedback. We also thank many people for the helpful discussion which have indirectly contributed to this paper, including: J. Butler (CERN, CMS), T. Camporesi (CERN, CMS), A. Charkiewicz (CERN, CMS), J. Closier (CERN, LHCb), A. Cook (CERN, Human Resources Departmernt), C. D'Ambrosio (CERN, LHCb), F. Dittus (CERN, Atlas), E. Elsen (CERN, Director for Research and Computing), L. Esteveny (CERN, Alumni network), N. Grub (CERN, LHCb), M. Kasemann (CERN, CMS), G. Passaleva (CERN, LHCb), A. Talesca (CERN, Alice). Last, but not least we thank all the survey's respondents. This paper was financially supported by a collaboration agreement between CERN and the University of Milan (ref. Addendum FCC-GOV-CC-0046, EDMS 1570377, KE3044/ATS).

1. Introduction

Research infrastructures (hereafter, RIs) are facilities, resources and services that foster knowledge and innovation in different fields (EC, 2019). They include major scientific equipment such as particle accelerators, outer space probes, archives of scientific data, computing systems and any other knowledge-based tools to achieve breakthrough research and innovation in science (ESFRI, 2018).

RIs are important human and social capital incubators (Giacomelli et al., 2017; Bianchin et al., 2019; Florio, 2019). They put together a wide and varied community of scientists from different fields and background, whose joint effort and cooperation is ultimately addressed to push forward the frontier of science, much more than traditional research in academia or firms' labs (Florio, 2019). They also frequently involve coalitions of universities and research laboratories from different countries, acting as a collective intelligence environment (ESFRI, 2019; Malone and Bernstein, 2015). For this reason, they are turning into important locus where bright minds concentrate and top-level human and social capital accumulates. Universities and firms themselves often take advantage of these collaborations as RIs are much better equipped and efficient than traditional small-scale laboratories. Indeed, nowadays, cutting-edge research requires investments in methods and instruments that exceed the capacity of single labs, especially when undertaking highly risky projects aimed at tackling global societal challenges such as health and demographic change, food security, clean and efficient energy (OECD, 2019; ESFRI, 2016). As a matter of fact, collaborative research in science, technology, engineering and mathematics facilitates researchers in most fields and contributes to knowledge, scientific productivity, economic and social outcomes (European Commission, 2020; Giffoni and Vignetti, 2019; Bozeman et al., 2015; Boyer and Robert, 2006; Lee and Bozeman, 2005).

In this perspective, many universities are increasingly interested in providing opportunities for their students to spend a period of study and/or work in these renowned international settings where they can develop a wide variety of skills and expand their network of contacts (Anderson et al., 2013a; 2013b).

In line with this idea, this paper examines the role of RIs as human and social capital breeding places, with a focus on early-career researchers (ECRs)[2]. Although the whole community of scientists and other professionals may extensively benefit from collaborations with RIs, this aspect is extremely important for ECRs as RIs represent unique and extraordinary learning and networking environment, whose effects may spread-out throughout their lifelong professional career. In a RI, they may carry out cutting edge research, closely working with experienced and renowned scientists in challenging projects, interact with different types of experts and develop a wide set of skills which are needed in the labour market. These skills do not only include the capability to carry out high level theoretical and empirical research, data analysis,

[2] ECRs include under-graduated, graduated students or post-graduate scholars up to 5 years post-PhD (https://www.ische.org/early-career-researchers/)

software and computing skills, among others, but also the ability to manage complex projects with tight deadlines, solve problems, work in a multicultural and interdisciplinary environment to mention a few (Giacomelli et al., 2017; Catalano et al., 2018).

Our analysis concentrates on the European Organization for Nuclear Research (CERN), which is a prime example of large-scale RI. Several studies have quantitatively and qualitatively discussed the positive effects of a period of stay at CERN on the professional careers of researchers (Bianchin et al., 2019; Anderson et al., 2013b; Camporesi, 2001), including higher salaries compared to peers without such an experience (Catalano et al., 2018, Camporesi et al., 2017; Florio et al., 2016). This paper validates and extends previous results looking at the mechanism through which such salary premium is generated. More specifically, we distinguish between the effect of acquiring new scientific and technical skills (human capital effect) from a social capital effect related to the economic return of social connections established at CERN and which can be useful for the future career. We also refine the expected level of salary premium earned by ECRs with a learning experience at CERN. The empirical strategy is based on three different sets of data (Creswell, 2003): an initial survey in the frame of CERN Alumni Network (which extends previous results collected first by Florio et al., 2016 and then enlarged by Camporesi et al., 2017), a survey involving team leaders of ECRs at CERN, and information on the salary expectations of ECRs in the labour market, based on secondary data sources.

The rest of the paper is structured as follows: Section 2 describes CERN as a human and social capital incubator; Section 3 introduces the method, analysis and results related to the three different datasets, while Section 4 concludes the paper.

2. CERN as a human and social capital incubator

Human capital can be generally defined as the set of knowledge, skills, competencies and attributes embodied in individuals (Keeley, 2007). Beside developed or innate abilities, two main components determine its accumulation: formal education and expertise acquired through work experience (Blundell et al., 1999; Mincer, 1974). Indeed, both education and work experience are important vehicles of skills and competencies generation which are ultimately rewarded in the labour market (Mincer, 1974; Card, 1999; Blöndal et al. 2002; Psacharopoulos and Patrinos, 2004; Montenegro and Patrinos, 2014).

Social capital is a more controversial and multidimensional concept. Here, we focus exclusively on the productive value of social connections, adopting the Bourdieu's perspective whose work highlights the payoff of establishing relations among people. Such 'individual' and 'personal' dimension of social capital has been often found to positively influence the labour market (Scrivens and Smith, 2013). For example, information acquired through social connections and the strength of these connections is important in job search and career advancement (Scrivens and Smith, 2013; Aguilera, 2002). Seibart et al. (2001) show the relevance of social capital to successful careers as it facilitates access to information, resources and career sponsorship. Wolff and Moser (2008) show that networking is related to concurrent salary and to the growth rate of salary over time. Matsunaga (2015) also shows that social capital is a driver of salary increasing wage in the same way of human capital.

Established in 1954 and located in Geneva, CERN carries out research in fundamental physics by providing particle accelerator facilities and promoting international collaborations to push the frontiers of science and technology.[3] CERN operates the Large Hadron Collider (LHC), the largest particle accelerator in the world. It is based in a 27-kilometre underground ring located between Switzerland and France. In this device, protons and atomic nuclei are accelerated and made to collide by electric fields to study the structure of matter and open a new frontier of research into the knowledge of the universe[4]. Collisions are analysed in particle detectors positioned along the LHC; the most important are ATLAS, CMS, ALICE, LHCb.[5]

The particle physics research activity at CERN requires the involvement of a vast community of scientists, universities, research institutes and firms from different countries, including bright young researchers. As reported in Florio et al. (2016), the intake of ECRs for the LHC experiments during the period 1993-2025 amounts to approximately 36,800. This figure comprises around 19,400 master and doctoral students and 17,400 postdoctoral researchers (not including participants in limited duration training courses). ECRs usually apply for a training programme launched at CERN under the supervision of a team leader, such as a senior scientist from their university or research institution in charge of supervising their work at CERN.[6] There, ECRs have the opportunity to spend a period of study and/or work in this vibrant and stimulating environment by gaining experiences with new technologies, having direct access to unique data, interacting with people from different cultures and knowledge domains.

[3] https://home.cern/about/who-we-are/our-mission

[4] https://home.cern/science/accelerators

[5] ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Solenoid), ALICE. (A Large Ion Collider Experiment), LHCb (Large Hadron Collider beauty). Other experiments are TOTEM (TOTal cross section, Elastic scattering and diffraction dissociation Measurement), LHCf (Large Hadron Collider forward)., MoEDAL (Monopole and Exotics Detector).

[6] The number of 36,800 by Florio et al. (2016) counts ECRs aged less 35 years and specifically: *CERN doctoral students*; *CERN technical students*; *CERN Fellows* and *CERN Users*. *CERN doctoral students* are PhDs enrolled in the CERN Doctoral Programme, where PhDs in Applied Physics, Engineering or Computing have the chance to work on their thesis while spending from 6 to 36 months at CERN. *CERN technical students* are undergraduate students (bachelor or master) in Applied Physics, Engineering or Computing enrolled in the CERN Technical Student Programme, where they are given the opportunity to spend a practical training period at CERN to complete their final project. The training period goes from a minimum of 4 to a maximum of 12 months. *CERN Fellowship* is targeted to graduates from a university or a technical institute, who can join CERN to carry out research in particle physics, or in other fields related to applied sciences, engineering or technical fields. The categories of *Fellowships* includes both *Junior Fellowship* for people with a BSc or MSc degree and no more than four years of experience after completing their highest diploma; and *Senior Fellowship* reserved to PhDs or people with at least four years' experience post-MSc and a maximum of 10 years' experience. CERN Fellows are hired as CERN staff and are generally offered an employment contract for between 6 months (minimum) up to a maximum of 36 months. Finally, *CERN Users* are CERN's guest scientists, technicians and engineers sent to CERN as members of a visiting research team to contribute to the upgrade or analysis of experiments under a memorandum of understanding with their home institution. Users are not part of the CERN staff, but are 'Associated Members' and they stay at CERN does not have any duration requirement. However, our elaboration on CERN personnel statistics from 1990 to 2019, indicate that their average duration at the laboratory is about 50 months. The threshold of 35 years was chosen as to focus on the human capital formation at the early years of the research career and avoid to count senior scientists, whose main formation is poorly (or not at all) linked to their experience at CERN. Accordingly, Users and Fellows aged more than 35 years were not considered. Similarly, partecipants to CERN schools or other very-short courses were not considered as well, because of very limited training period at CERN. The same approach was followed in this paper. For details on CERN training programmes see https://careers.cern/alljobs, last access on 03.01.2021

Moreover, they can participate in meetings ranging from technical to managerial levels, actively contributing to conferences and workshops with established professional and renowned scientists. Such a unique experience allows them to develop skills, competences and networks of contacts which are attractive and useful in almost all workplaces also outside research, including industry and the financial sectors (Bianchin et al., 2019; OECD, 2014; 2019; Boisot et al., 2011; Camporesi, 2001).

Few authors have already reported on the positive effect of CERN concerning the creation of human capital and how skills acquired in this environment are highly valued both in the academic and non-academic field (Laurila, 2013; Anderson et al., 2013b; Danielsson, 2013). However, the entity of such human capital accumulation, including its quantification, for example in terms of salary premium, and the mechanism through which such premium is generated, including the role of social capital, have been underplayed. For example, Camporesi (2001) carried out a survey with 671 students involved at CERN's *Large Electron Positron Collider* (LEP)[7]. Most of these students acknowledged that their activity in the LEP project has undoubtedly played a significant role in their professional career. The study shows that the private sector was explicitly interested in the skills acquired by students during their learning experience at CERN, such as the ability to work effectively in large and diverse teams and to solve complex problems, the exposure to cutting edge technologies in the electronics and computing domain, the familiarity with software techniques related to handling large quantities of data and performing sophisticated modelling.

In a similar vein, Anderson et al. (2013a) reported findings from a survey to ECRs in High Energy Physics (HEP) in the United States to record attitudes and interests, including their career outlook. The authors collected a total of 1,112 responses. Around 30% of the total respondents were at CERN, 40% in universities, and the remainder in other research centres. Programming, data and statistical analysis as long with communication and writing skills are considered the most valuable skills learned in HEP. These skills are valuable for respondents with professional careers outside academia (around 6%) as well. Anderson et al. (2013b) also provide two examples of students previously working at CERN (both for the ATLAS collaboration) who then moved to other sectors and explained how a learning and experiential process in RIs was instrumental in creating skills needed outside the basic research field.

Lately, Giacomelli et al. (2017) and Bianchin et al. (2019) reported on a recent survey conducted by the CERN Alumni initiative[8] to investigate the value of the education and skills acquired at CERN. Building on a pilot survey, a second version of the questionnaire (targeting theoretical physicists and investigating on the role played by CERN for this community) focused on both current and past members

[7] LEP collider was – and still is – the largest electron-positron accelerator ever built. It closed down on 2 November 2000 to make way for the construction of the Large Hadron Collider in the same tunnel. (https://home.cern/science/accelerators/large-electron-positron-collider)

[8] https://alumni.web.cern.ch, accessed on July 2, 2018.

of LHC and other CERN experiments. 2,692 responses were collected.[9] The survey revealed that around one-third of respondents moved from HEP to other sectors of work. Out of these people, 70% of respondents consider the experience at CERN positive or very positive in obtaining their current job. The majority of people leaving CERN currently works in the private sector, specifically in ICT, engineering, consulting and other domains, where they are employed as managers, directors and in other executive-level positions. Surveyed people declared to have acquired a variety of skills at CERN which are considered to be important in their current job, and specifically related to programming (465 respondents), working in an international team (459), data analysis (450), logical thinking (389) and communication (381).

Florio et al. (2016) present the first attempt to calculate the salary premium arising from a training experience at CERN LHC as part of a wider cost-benefit analysis of the research infrastructure. The authors carried out a survey with 384 current and former LHC ECRs from more than 50 countries with the objective to measure the impact of their period of stay at CERN on their expected salaries. The authors estimated an average expected salary premium of 11.8%, meaning that, on average, each researcher at CERN LHC would expect to earn every year, and along with his/her long life career, a salary 12% higher than a peer without such experiential learning at CERN.

A more detailed statistical analysis was carried out by Camporesi et al. (2017) with the objective to study the key determinants of the salary premium. By using ordered logistic regressions, Camporesi et al. (2017) came up with three main results. Firstly, they confirm the existence of an expected CERN salary premium in the range of 5-12%; secondly, they found that there is no statistical difference between the expected salary premium declared by current CERN ECRs and the salary premium of former CERN ECRs in the labour market. Thirdly, the drivers of the expected salary premium were the duration of the stay at CERN and the type of skills acquired during the stay. The salary premium mainly depends on having acquired technical skills (e.g. knowledge of programme languages, mechanical equipment, tools, etc.).

This paper goes in this direction. We want to expand these previous analyses (by extending the survey carried out by Florio et al., 2016 and Camporesi et al., 2017) on the expected salary premium and further investigate on learning-by-practical-experience at CERN as a driver for future professional opportunities accruing to ECRs. In particular, we want to refine the level of the expected salary premium, looking and comparing different sources of data and investigate the role of social capital besides the acquisition of skills as a primary driver of successful future careers.

[9] 97% of respondents were from LHC experiments and 3% from other CERN experiments,.

3. Methods, design of the study, and implementation

We propose a methodology framework which mixes quantitative analysis and qualitative evidence coming from three different datasets: i) data from an online survey to the CERN Alumni Network (which extends previous surveys carried out first by Florio et al., 2016 and Camporesi et al., 2017); ii) data from an online survey to Team Leaders iii) a database on ECRs' salaries we assembled by exploiting secondary data sources.

3.1 Survey to Early-Career researchers of CERN (Alumni Network)

Data

The first pillar of our empirical framework is a survey targeted to ECRs at CERN which includes both ECRs at the time of the survey (hereafter *current ECRs*), and ECRs who have completed their period at CERN and are now in the job market (*former* ECRs). Following Camporesi et al. (2017), the survey was organised into four sections. Section A refers to personal information of ECRs including their gender, nationality and educational background, among others; section B refers to their experience at LHC including carried out activities and developed skills. Section C investigates career expectations of current ECRs, while Section D investigated the current situation and future career aspiration of former ECRs.

The survey was launched during the "CERN Alumni first collision event" in February 2018 and also disseminated by CERN through its Alumni platform which counts about 3,000 registred members. In total, we collected 438 valid questionnaires with a response rate of 15%.[10]

Respondents came from more than 50 countries, mostly from CERN member states (65%)[11]. 75% of participants were male, while 60% had already left CERN and, at the time of the survey, were employed elsewhere in the labour market (*former ECRs*). 57% of the respondents owned or were completing a PhD, while the remaining share held either a bachelor or a master degree. Most of the respondents (83%) had a background in theoretical, applied or experimental physics; the remaining 17% mostly included computer scientists, engineers and a small portion of mathematicians and administrative staff. The average age of participants was 35 years (31 years for current ECRs and 37 for former ECRs). The average duration of stay at CERN was around 44 months (3.7 years).

Table 1 shows that, during the period at CERN, the great majority of respondents acquired technical (86%) and scientific skills (85%), followed by problem-solving capacity (77%). Developing and using networks of collaborations and developing team or project leadership ranked last. Hence, according to

[10] The risk of self-selection - for which only satisfied alumni would have been more prone to participating in the survey – could be partly mitigated by the large number of face to face interviews (50%) conducted directly at CERN. We interviewed ECRs regardless of their degree of satisfaction with CERN activities. Several reminders to solicit answers to the survey have also been sent by e-mail to all target group.

[11] Austra, Belgium, Bulgaria, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherland, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Spain, Sweden, Switzerland, UK.

the majority of respondents, the experience at CERN was important, primarily, to increase technical and scientific competences; in contrast, social connections and soft skills such as communication and creativity or networking opportunities are considered of secondary importance. This result is in line with their decision to apply for a research period at CERN where most of the respondents (90%) selected the option ‘Deepening the knowledge and competences in the scientific domain of interest’, followed by the option ‘Develop new professional skills’ (82%). The options ‘Opportunity to work with world-class physicists’ and ‘World undisputed prestige of CERN’ were ticked by a lower percentage of participants (71% and 58% respectively).

<TABLE 1>

Respondents also revealed their actual and expected start and end-career salaries. Regardless of being current or former ECR, about 70% of respondents indicated start-career salaries in the lowest salary categories (less than EUR 50,000) (Figure 1). In contrast, answers related to end-career salaries concentrate on the highest categories (at least EUR 50,000) for both current (87%) and former ECRs (91%). Additionally, there is a statistically significant difference in the distribution of expected salaries between those stated by current and former ECRs. Current ECRs are more optimistic than former ECRs as regards initial salaries and more pessimistic as concerned end-salary expectations, pointing to an underestimation of salary levels in the long-run by current researchers.[12] We take on board this difference in the econometric analysis.

<FIGURE 1>

Furthermore, the great majority of respondents believe that their salary levels are (will be) higher than the salary earned by their peers without the experiential learning at CERN. 34% (124) of respondents ticked the option zero (no difference exists), while the remaining 66% choose either the option 'up to 10%' (29% of respondents) or 'from 11 to 30%' (25% of respondents) or the option 'more than 30%' (12% of respondents). Lastly, most respondents would see themselves in the future working in research outside academia (205 ticks), followed by academia (157 ticks), ICT and finance (74), industry (73), and public sector (18 ticks).

[12] The statistical difference was tested by applying the chi-square test (χ^2 =10.5; p-value<0.05) to the distribution of start career and the Fisher’s exact test (p-value<0.05) to the distribution of end-career salary.

3.2 Econometric analysis: model and results (human capital vs social capital) and salary premium

The objective of the econometric analysis is twofold. First, we want to understand the main factors driving ECRs' salary expectations with a focus on the effect of CERN. By controlling for personal characteristics, we try to distinguish between the effect deriving from the acquisition of scientific and technical skills at CERN (*Human capital effect*) and the role of new social connections and networking (*Social capital effect*). Second, we want to refine the level of expected salary premium estimated in Camporesi et al. (2017) by exploiting the fresh evidence gathered with the new survey.

In a cost-benefit analysis perspective, benefits accruing from training activities, with an impact on human capital accumulation, are measured by considering the lifelong salary. Therefore if a premium would exist, it should be investigated by looking at the end-career salary levels (Florio, 2019; Florio et al., 2016; European Commission 2014, ch. 7). We report in the appendix further analyses and robustness checks where we consider start-salary as a variable of interest.

We use an ordered logit model where the dependent variable is an observed ordinal variable y_i which reflects different level categories of end-career expected salary. The dependent variable takes integer values from 1 to J, where i identifies the respondent. The ordinal salary response y_i with J categories can be represented as an underlying (latent) continuous response y_i^* such that:[13]

$$\begin{aligned} & y_i = 1 \text{ } if \text{ } y_i^* \leq \alpha_1 \\ & y_i = j \text{ } if \text{ } \alpha_{j-1} < y_i^* \leq \alpha_j \quad j = 2, \dots, J-1 \\ & y_i = J \text{ } if \text{ } \alpha_{J-1} < y_i^* < \infty \end{aligned} \tag{1}$$

where the parameters α_j, called cut-points, are free parameters to be estimated and are expected to be in increasing order ($\alpha_1 < \alpha_2 < \cdots < \alpha_{J-1}$). Given a set of covariates denoted by the vector X_i, which in our case includes the demographic traits of the respondent, the length of stay at CERN, the acquired human and social capital among others (see Table 2 for the full list of variables), the conditional distribution of y_i given X_i is represented by the function:

$$\Pr(y_i = j|X_i) = \Lambda(\alpha_j - X_i\beta) - \Lambda(\alpha_{j-1} - X_i\beta); \quad j = 1, \dots, J \tag{2}$$

In Eq. 2, $\Lambda(.)$ denotes the logistic cumulative distribution function and allows estimating the probability that the respondent selects one of the submitted salary categories given the set of regressors X_i. The effect, if any, of human and social capital on the (expected) stated salary levels is investigated by

[13] The ordered logit model in Eq, 2 is equivalent to a system of equations consisting of a set of cut-points α_j and a liner regression for the non-observed latent continuous variable y_i^* (i.e. the exact level of salary within the proposed salary catagories to the respondent) such that: $y_i^* = X_i\beta^* + \varepsilon_i$, where ε_i is an error term that follows a logistic distribution with mean zero and standard deviation σ_ε^* (Balakrishnan 1992). For more details on the link between this linear model and the model in Eq. 2 see Grilli Rampichini (2014).

calculating the marginal effects of such variables on the response variable y_i. The alpha and beta coefficients in Eq. 2 are estimated through maximum likelihood (Long and Freese 2014).

The vector of the slopes β is not indexed to the salary category j, therefore the effects of regressors are the same across the salary categories. This feature of the model is known as parallel regression (or proportion odds) assumption. We test the parallel regression assumption for each specification of Eq. 2. Table 2 illustrates the main variables (some of them were pre-treated) while table A1 in the appendix reports their statistical distribution.

<TABLE 2 >

Table 3 shows estimates of seven different specifications, one in each column. We are particularly interested in the regressors *Scientific human capital*, *Social capital (networking)*, *Other soft skills,* and *Length of stay*, whose coefficients are expected to be positive and statistically significant. Whatever the specification, we always control for the experiment where the respondent was working during her stay at CERN (*Experiment fixed-effects)* and her nationality (*Nationality fixed-effects)* as salaries expectations are likely to depend on national labour market conditions and educational systems (Bianchin et al., 2019). On top of that, we also control for whether the interview was carried out *Face-to-face* or was not to take into account different respondent's behaviour caused by the different interview modality. Lastly, for each specification, the test (p-value) for the proportional odds assumption is reported (Long and Freese 2014, ch. 7). As shown in the last row of the Table 3, it is met in all specifications.

The first specification (Column 1), focuses on the role played by the *Scientific human capital* on salary expectations. Its positive coefficient means that the likelihood of declaring higher salaries increases with scientific and technical skills acquired during the stay at CERN. In contrast, the coefficient for *Social capital (networking)* and that for *Other soft skills* lack of statistical significance both when are considered alone (Columns 2) and when are analysed together with the *Scientific human capital* (Column 3). This indicates that developing networks and soft skills such as communication skills, independent thinking and creativity are likely not to affect expected salaries in line with previous studies on the career trajectories of people with working experience at CERN (Bianchin et al., 2019; Giacomelli et al., 2017). Indeed, scientific skills such as data analysis and programming skills acquired at CERN are considered extremely important for ECRs' future career. Additionally, while most respondents will continue their careers as software engineers, researchers and data analysts, they consider networking opportunity at CERN less effective to find a job outside HEP (Bianchin et al., 2019; Anderson et al., 2013a).

In the specification reported in Column 4, we added the variable *Length of stay* which shows a positive and statistically significant coefficient at 10% level. The longer the duration of stay at CERN, the higher is the probability that respondents state higher salary levels, once the type of skills acquired has been accounted for. To further investigate on this issue, we allowed the *Length of stay* to interact with *Scientific human capital*. This is done in Column 5, which shows that the skills acquired are likely to increase with the duration of stay and this, in turn, positively correlates with expected salaries.[14]

The specification in Column 6 introduces our full model, where personal characteristics (*Male*, *Age*, *PhD*), and career-related information such as *Salary for comparators*, and sectors (*Research and Industry_Finance*) are jointly plugged into the same model. By including these variables, the *Scientific human capital* and the *Length of stay* at CERN keep their predictive power in explaining salaries. As expected, men expect higher salaries with respect to women as HEP is a male-dominated domain (Camporesi et al. 2017; IOP Institute of Physics, 2012). The sector where ECRs currently work or expect to continue their careers also highly influences salary expectations. Specifically, ECRs in the *Research* sector, either in or outside academia have lower salary expectations as compared to people in industry or finance. *Salary for comparators* is a variable capturing to what extent respondents believed that their (expected) salaries are higher than the ones of their peers. While this variables should not be intended as a substitute for actual salaries arising from, e.g. a control group based on a counterfactual approach,[15] it has been shown to be a key driver of the likelihood of declaring higher salaries both in HEP (Camporesi et at., 2017) and other different context (Schweitzer et al. 2014). Our analysis confirms that the higher the salary respondents expect to earn with respect to their peers thanks to their learning experience at CERN, the higher the probability that they point to higher salaries categories is.

One may argue that the specification in Column 6 is not sufficiently reliable as a two-way (causal) relantionship may exist between the variable *Salary for comparators* and our dependent variable *End-career salary expectation.*[16] Such reverse causation is likely to lead to an overestimation of the coefficients reported in Column 6, and especially the one for the *Length of stay,* which, in turn, is used to calculate the salary premium (see below). We address this concern in Column 7, where we re-estimate the model proposed in Column 6, without the variable *Salary for comparators.* The new specification returns very similar results as the previous one, and more importantly, the chi-square test on the difference between the coefficient on the *Length of stay* in the two specifications does not reject the null

[14] The negative sign and the absence of statistical significance of the coefficients for "*Scientific human capital*" and "*Length of stay*" is due to the statistical correlation between that variables and the interaction term "*Length of stay*Scientific human capital*". The correlation coefficients are equal to 0.56 and 0.88 respectively and both statitistically significant at 1% level. We also analised the interaction between the "*Length of stay*" and "*Social capital (networking)*", but the associated coefficient is not statistical significant at all.

[15] A control group would be a group of ECRs with similar personal traits and educational careers to our sample of ECRs, but without the training experience at CERN.

[16] We thank an anonymous refree for this suggestion.

hypothesis of the statistical equivalence between the two coefficients ($\chi^2 = 0.03; p - value = 0.8536$). This ensures that the salary premium can be calculated either from the specification in Column 6 or from the specification in Column 7 without significant statistical differences.

Moreover, the variable *Being former ECRs* loses its predictive power in explaining end-career salaries in specifications when the duration of stay at CERN, personal traits and career-related information are considered (Columns 4, 6, and 7). This corroborates the idea that salaries trajectories are shaped by personal characteristics of ECRs, work experiences and sector of employment regardless of whether being current or former ECRs at CERN.

Finally, the *Likelihood ratio test* indicates that, in each specification, regressors have significant effects on our dependent variable, while the *Count* R^2 indicates that 75% of respondents' records are correctly predicted.

It should be noted that one of the caveats of the econometric analysis is linked to the number of observations entering in the model specifications. They go from 345 to 320 as compared to a number of 367 observations valid for the expected salaries in Figure 1. That happens because respondents skipped some questions while answering the questionnaire. While causing potential distrorsion in the results, the econometric analysis is just one of the pillars our research is based on, therefore we argue that the possible bias stemming from the respondents behaviour is partly mitigated by the evidence from the other two pillars, i.e. the survey to the Team Leaders and the analysis of the secondary data.

<TABLE 3>

As a second step of this econometric analysis, we follow Camporesi et al. (2017) to estimate the expected salary premium by using Eq. 2. The salary premium is calculated by analysing the relationship between the length of the experiential learning at CERN and the stated end-career salary level revealed by respondents considering the full model (Column 6 in Table 3).[17] Table 4 reports the marginal effects of the *Length of stay* on salary levels, i.e. the predicted probabilities of one additional month of stay at CERN on the five outcomes of our dependent variable. Marginal effects are expressed in percentage terms.

The negative coefficients for *Length of stay* on the three lowest salary categories (< 30,000 EUR, 30,000-40,000 EUR; 40,000-50,000 EUR) mean that the longer the duration of the experiential learning at CERN the lower the probability that respondents' state a low salary category; in contrast, the length

[17] For the sake of transparency, we also calculated the marginal effects associated with the variable *Length of stay* from Column 7 in Table 3. As expected, they are not statistically different from the marginal effects reported in Table 4. Starting from the lowest (< 30,000 EUR) to the highest salary category (> 60,000 EUR), they are as follows: -0.057* (0.030); -0.034* (0.020); -0.040* (0.021); 0.118** (0.055); 0.250*** (0.105). Numbers in paratheses denote the standard devation, while asterisks denote statistical significance at the 1%, 5% 10% level respectively as usual.

of stay increases the likelihood of declaring salaries in the highest categories, i.e. between EUR 50,000 and 60,000, and more than EUR 60,000. Specifically, an additional month of experiential learning at CERN increases the probability of stating a salary between EUR 50,000 and 60,000 by 0.12 percentage points ($p - value < 0.05$) and a salary more than by EUR 60,000 by 0.25 percentage points ($p - value < 0.05$), ceteris paribus.

The average duration of stay at CERN in our sample was about 44 months, and specifically, 24 months for current ECRs and 60 months for former ECRs. Accordingly, for an 'average' respondent the expected salary premium would amount to 5.3% (about 3% for current ECRs and 7% for former ECRs) if the salary category is EUR 50,000- 60,000 and to 11% (6% for current ECRs and 15% for former ECRs) if the salary category is higher than EUR 60,000. This is our best guess for the value of salary premium as perceived by researchers who have spent a period of their study at CERN, calculated by taking on board an information set including personal traits and career-related information. The estimation is in line with previous works which see CERN as a human capital incubator.

<TABLE 4>

To sum up, the survey to CERN researchers shows that, after controlling for personal characteristics and other factors, the likelihood of expecting higher end-career salaries increases with the scientific and technical skills acquired during the period at CERN. Such skills are likely to increase with the duration of stay at the RIs and translate into an expected average salary premium ranging from 5% to 11%. In contrast, the possibility of developing networks and other soft skills, although beneficial for future careers, do not affect ECRs' salaries expectations.

3.3 Survey to Team Leaders of Early-Career Researchers

The second pillar of our empirical framework is a survey to Team Leaders which is a novelty of our research compared to previous studies. Team Leaders are senior scientists affiliated at universities or other research institutes that are instrumental for ECRs to start their period of study and/or work at CERN. Their role is to supervise the research activity of ECRs. They provide additional valuable information and comparisons about the career outcomes of students who had the opportunity to spend a period at CERN with other students who did not have such a learning experience. This survey was structured into five questions and investigated different issues such as: i) the share of supervised students spending a research period at CERN as compared to those applying for alternative experiences or with no external experience; ii) the perception of team leaders about the reasons why students decided to spend a research period at CERN; iii) the contribution of CERN in improving ECR's skills, according to their perceptions; iv) their opinion on the reliability of salary expectations stated by ECRs reported in

Camporesi et al. (2017); v) team leaders' expectations about the number of future ECRs at CERN.

The survey to team leaders was launched in parallel to the ECRs survey (from March to May 2018) and 322 valid responses were collected, representing a response rate of about 29% of the potential target. At the time of the survey, 38% of team leaders were involved in ATLAS, 49% in CMS, 34% in the LHCb and the remaining share in other CERN experiments. On average, 38% of the ECRs - supervised by team leaders – have spent one year or more at CERN, another 37% less than one year, while the remaining percentage have spent a research period elsewhere or did not spend any period of research outside the university.

Concerning the main motivations for applying for a research period at CERN, the majority of team leaders rated as important or very important the following options: deepening the knowledge and competences in the domain of work (92%), working in an international environment (92%), the possibility to work with world-class scientists (93%) and developing new professional skills (90%) (Table 5).

<TABLE 5>

Furthermore, Table 6 shows that, according to team leaders, the most important skills acquired by researchers who have spent a period of work and/or study at CERN are soft skills such as networking (85% of respondents rated the contribution of CERN as 'to a high extent' and 'very great extent'), followed by communication (76%), scientific skills (71%) and technical skills (70%).

<TABLE 6>

Hence, these responses highlight that, according to team leaders, spending a period at CERN is not only beneficial to deepen knowledge and competences and to improve scientific and technical skills but also to develop useful connections which would be instrumental for ECRs' future careers. Although, this finding may appear partially in contrast with the responses given by ECRs in the previous survey and with the econometric analysis, the survey to team leaders gives a wider perspective to our study, which is difficult to be captured by ECRs themselves and quantitative tools. Team leaders' opinions are based on their tenured experiences and long-term views. They are able to see and understand the 'value' that networking activities stemming from the varied and multicultural environment of CERN can offer on career trajectories, for example in terms of access to information, reputation, resources and career sponsorship.

Concerning the salary premium, 85% of the surveyed team leaders confirm the findings on the salary premium reported in our present survey to the ECRs and previous surveys (Camporesi et al., 2017) (Table 7). Specifically, half of them (54%) stated that a premium between a minimum of 5% to maximum of 12% is reasonable, while one third (31%) would have expected an even higher impact. Only 3% of them would has expected a lower salary premium. The remaining share of 12% did not have sufficient information to give an opinion. This points to a consensus about the level of salary premium amongst HEP research community since team leaders know different cohorts of students and also students not spending any time at CERN, which can be considered a counterfactual sample.

<TABLE 7>

When asked about the future number of university students that would apply for a training period at CERN, the majority of team leaders (53% of respondents) suggest that the number of students applying at CERN will likely remain stable, followed by 20% declaring that it will likely increase if new projects at CERN will start. Overall, these findings suggest that the role of CERN as human and social capital incubator is likely to become even more important in the next future.

3.3 Analysis of secondary data

In addition to the two surveys described above, we explored secondary data sources and created a second salary database with the objective of comparing our estimations of the salary premium with additional information on the salary expectations of ECRs in the labour market. More specifically, our analysis relies on the assumption that a research period at CERN can be compared with a year of doctoral studies spent at universities elsewhere. In both cases, students may acquire scientific and technical expertise and other skills - such as understanding and analysing a large amount of information, managing complex projects, preparing concise written documents, attending conferences, and working under pressure and tight deadlines - that are relevant in their careers (in terms of effects on the salary premium) and which cannot be gained by their peers without these experiences.

To estimate the salary premium of doctoral studies, we collected information on the salary differences between people with a master degree and those with a doctoral degree from national and international statistics. Since the main area of reference is HEP we particularly concentrated on people with PhD in physics (Pold and Mulvey 2015; Mulvey and Pold, 2016, 2017). In this subject, data are only available for the US labour market.

Early-career salaries for people with a doctorate in physics (cohort 2013-2014) have a median of USD

71,000 per year (Mulvey and Pold, 2016)[18]. The equivalent salary after one year for those with a master's degree was USD 53,000 (in the period 2012-2014) [19] (Pold and Mulvey, 2015). Although the samples are small, reported data would imply a 'premium' of around USD 18,000 (34%) at the starting level in the US. Assuming that the average number of years needed to earn a doctoral degree is four years[20] – a single year of doctoral study corresponds to a salary premium of USD 4,500 (8.4%) which is in the range (5%-11%) found in our analysis. Lower returns for doctoral studies are found by the NACE (National Association of Colleges and Employers) salary survey in other scientific sectors.[21] Results show a total doctoral study premium of 24% in math and science, 28% in engineering and 37% in computer science. A single year of doctoral studies would correspond to a premium of 5.9%, 7.0% and 9.2% respectively. All these percentages are comparable to the potential CERN salary premium estimated above.

In continental Europe and for natural sciences in general, the doctorate premium is about 20%,[22] which means around 4.3% per year. It is a lower percentage compared to the US one and more consistent with the premium estimated in our analysis. This implies that, on average, spending a period of study and/work at CERN rewards as much as spending one year of doctoral studies in Europe in scientific subjects.
Hence, our assumption suggests that ECRs with an additional doctoral year of study (and no experience at CERN) typically enjoy an extra 'premium' compared to those without such experience (e.g. master students). Here, we also assume that ECRs (either with a master or a PhD) with the experience at CERN get a salary premium of a similar entity of that of a student with an extra year of doctoral study compared to another student without such extra year of doctorate. However, in our sample, there is no statistical difference in the premium expected by master and PhD graduates with an experience at CERN (see table 3). This would imply that (i) the PhD at CERN would get an additional premium on the top of the premium they would have earned in a scenario of 'no experience at CERN' as compared to master students. (ii) the premium related to the experience at CERN would be higher in the case of a master student since his/her salary is lower than a PhD graduate without an experience at CERN. Hence master students would particularly benefit from the experience at CERN, because without such experience their expected salary, on average, would have been, even lower than a PhD graduate without the experience at CERN.

Another useful comparison between primary and secondary data is to consider the salary difference in terms of the skill acquired during the period spent at CERN. A detailed survey of the economic value of

[18] USD 99,000 permanent full-time jobs in the private sector - 31% of respondents; USD 66,000 for those working in government positions - 14% of respondents; 48,000 for those working in academia – 52% of respondents.

[19] USD 65,000 in the private sector – 53% of respondents; USD 41.000 in the academia – 19% of respondents. https://www.aip.org/sites/default/files/statistics/employment/ms1yrafterdeg-p-14.pdf, accessed on March 7, 2018.

[20] An examination of different PhD courses respectively in Europe and US suggested an average length of around 4 years

[21] https://www.tougaloo.edu/sites/default/files/page-files/2017-nace-salary-survey-winter.pdf, accessed on March 7, 2018.

These figures refer to base salaries only, hence they do not include bonuses, commissions, fringe benefits or overtime rates. Data are obtained by surveying NACE employer members from August 2016 to November 2016 for a total of 243 people.

[22] https://www.academics.com accessed on March 7, 2018. The average yearly salary for a general master's degree is EUR 47,200 per year. For doctoral degrees, it is on average EUR 55,266 per year.

specific skills remains yet to be carried out; however, some initial indications exist. According to the PayScale survey data,[23] a physicist with data analysis skills has an average salary of USD 93,140, while the average salary for a research position in physics is USD 89,000 and USD 75,000 for a physicist with generic skills (see table A5 in the appendix). The total yearly salary premium for a physicist with specific skills is in the range of 5.5% to 26.7% as compared to a physicist with generic skills. Based on these differences, the skills acquired by people who spend a research period at CERN can, therefore, be considered for an additional salary premium compared to people who have not carried out a significant period of applied research and development during their studies. This result stresses on the importance that the human capital effect, measured by the acquisition of scientific and technical skills, plays in the generation of the salary premium as compared to other impacts related, for example, to the role of social capital.

<TABLE 8>

[23] https://www.payscale.com/research/US/Job=Physicist/Skill accessed on March7, 2018
Founded in 2012, PayScale helps people and businesses to obtain accurate, real-time information on job market compensation.

4. Conclusions

RIs are large-scale collective scientific enterprises which foster research, innovation and knowledge-sharing. They are important human and social capital incubators where bright minds concentrate and top-level human and social capital accumulates. The case study adopted in this paper refers to a world-famous European RI which hosts the largest accelerator used for research in particle physics: CERN. This RI is organised in the form of a large collaboration among different countries and involves a large number of universities and research institutes.

In a cost-benefit analysis framework, the contribution to human and social capital formation is one of the most important socio-economic benefits of a RI, in particular for early-career researchers who have the opportunity to develop competencies, skills and social connections in a unique learning environment. Based on the collection of primary and secondary data, we found that spending a period of study and/or work at CERN (particularly at LHC programme and experiments) is particularly instrumental in developing scientific and technical skills which are attractive in the job market both inside and outside the academia. Such skills, according to ECRs perceptions, would translate into a salary premium ranging between 5% and 11% as compared with a scenario where they do not carry out such a unique experience. The total lifetime salary 'premium' also varies depending on personal characteristics, future sector of activity and the total duration of stay.

Developing useful social connections is also important for job search and career advancement and may also contribute to this 'premium'. Although this aspect was weakly perceived by the ECRs , its importance was particularly stressed by team leaders. The latter have a wider perspective on the future career of a different cohort of students and are more likely to understand the 'value' of networking activities based on their tenured experiences and long-term views compared to ECRs. The 'individual' dimension of social capital, which mainly acts through social connections established by ECRs working at the RI and the reputation of the infrastructure can also influence the labour market positively, for example, facilitating access to information, resources and career sponsorship.

We also found that the salary premium of spending a period at CERN can be ultimately compared with a year of doctoral studies in Europe. In both cases, students acquire scientific and technical expertise and other skills that are relevant for their careers and can build connections with other researchers and professionals in academia and in the industry. However, the acquisition of scientific and technical skills seems again to play a more significant role in the generation of the salary premium.

Initial findings reported by earlier studies have been confirmed and expanded by the new evidence collected with complementary methods. Since the employers value the additional skills and experience, training at CERN generates a measurable socio-economic added value and this is something that governments should also positively consider when deciding to invest public money in RIs. This methodology can be applied in the future to estimate the socio-economic return of other RIs in human and social capital accumulation. Estimates of the expected salary premium at CERN can be used as a benchmark for future cost-benefit analysis of post-LHC particle-collider research infrastructure. Future research on this topic should also consider that students that are given the opportunity to train in prestigious work environments like CERN tend to be the most prepared and talented ones, also due to rigid selection processes at different levels. Hence, the salary effect discussed in this paper may partly stem from this initial bias, something that should be better considered in a future survey and research design by resorting, for instance, to techniques based on the counterfactual approach (see among others, Angrist and Pischke, 2008). Indeed, the comparison between our ECRs with a CERN experience and "similar" ECRs who did not have the same experience would minimise further bias generated by (self-) selection.

References

Aguilera, M. B. (2002). The impact of social capital on labor force participation: Evidence from the 2000 Social Capital Benchmark Survey. *Social science quarterly*, *83*(3), 853-874.

Anderson, J., Asaadi, J., Carls, B., Cotta, R., Guenette, R., Kiburg, B., & Penning, B. (2013a). Snowmass 2013 *Young Physicists Science and Career Survey Report*. arXiv preprint arXiv:1307.8080.

Anderson, J., Brock, R., Gershtein, Y., Hadley, N., Harrison, M., Narain, M., ... & Prebys, E. (2013b). *Benefits to the US from physicists working at accelerators overseas*. arXiv preprint arXiv:1312.4884.

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Bastianin, A., & Florio, M. (2018). *Social cost-benefit analysis of HL-LHC. CERN-ACC-2018-0014*, https://cds.cern.ch/record/2319300/files/CERN-ACC-2018-0014.pdf

Bianchin, C., Giacomelli, P., Iconomidou-Fayard, L., Niedziela, J., & Sciascia, B. (2019). *Study on the career trajectories of people with a working experience at CERN*. arXiv preprint arXiv:1910.05116.

Blöndal, S., Field, S., & Girouard, N. (2002). *Investment in human capital through upper-secondary and tertiary education*. OECD Economic Studies, 2002(1), 41-89.

Blundell, R., Dearden, L., Meghir, C., & Sianesi, B. (1999). *Human capital investment: the returns from education and training to the individual, the firm and the economy*. Fiscal studies, 20(1), 1-23.

Boisot, M., Nordberg M., Yami S., & Nicquevert B. (2011). *Collisions and Collaboration: The Organization of Learning in the ATLAS Experiment at the LHC*. Oxford University Press.

Boyer, M. and J. Robert (2006). *The Economics of Free and Open Source Software: Contributions to a Government Policy on Open Source Software*. CIRANO Project Reports 2006rp-03, CIRANO.

Bozeman, B., Gaughan, M., Youtie, J., Slade, C. P., & Rimes, H. (2015). *Research collaboration experiences, good and bad: Dispatches from the front lines*. Science and Public Policy, 43(2), 226-244.

Camporesi, T. (2001). *High-energy physics as a career springboard*, European Journal of Physics, 22, 2, 159-148.

Camporesi, T., Catalano G., Florio M., & Giffoni F. (2017). *Experiential learning in high energy physics: a survey of students at the LHC*. European Journal of Physics, 38(2), 025703.

Camporesi, T., Catalano, G., Florio, M., & Giffoni, F. (2016). *A" LHC Premium" for Early Career Researchers? Perceptions from within*. arXiv preprint arXiv:1607.01941.

Catalano, G., Florio, M., Morretta, V., & Portaluri, T. The Value of Human Capital Formation at CERN CERN-ACC-2018-0025 (CERN, Geneva, Switzerland, August 22, 2018). *Preprint CERN-ACC-2018-0025, https://cds. cern. ch/record/2635864*.

Catalano, G., Del Bo, C., & Florio, M. (2015). *Human Capital Formation at LHC: Survey Results*. Available at SSRN 2722467.

Card, D. (1999). The causal effect of education on earnings. In *Handbook of labor economics* (Vol. 3, pp. 1801-1863). Elsevier.

Creswell, J. W. (2003). '*Research design: Qualitative, quantitative, and mixed methods approach',* Sage publications.

Danielsson H. (2013). Students as a Bridge Building Link. CERN Presentation at NORDTEC-Conference in Lund, June.

ESFRI (2016). *Strategy report on research infrastructure, Roadmap 2016*, available at http://www.esfri.eu/sites/default/files/20160308_ROADMAP_single_page_LIGHT.pdf

ESFRI (2018). *Roadmap and strategy report on research infrastructure*' http://roadmap2018.esfri.eu/media/1066/esfri-roadmap-2018.pdf

ESFRI (2019). *Guidelines on cost estimation of research infrastructures*. Report developed for StR-ESFRI (Support to Reinforce the European Strategy Forum on Research Infrastructures) by CSIL – Centre of Industrial Studies in Milan.

European Commission (2014). *Guide to Cost-Benefit Analysis of Investment Projects Economic appraisal tool for Cohesion Policy 2014-2020* https://ec.europa.eu/inea/sites/inea/files/cba_guide_cohesion_policy.pdf

European Commission (2019). *European Research Infrastructure 'What Research Infrastructures are, what the Commission is doing, strategy areas, funding and news'* https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures_en

European Commission (2020). *Designing a socio-economic impact Framework for research infrastructures:*
preliminary lessons from the RI-PATHS project. In Beck, H. & Charitos, P (Eds.) Economics of Big Science. Science Policy Reports. ISBN 978-3-030-52391-6

Florio, M. (2019). Investing in science: social cost-benefit analysis of research infrastructures. MIT Press.

Florio, M., Forte S., and Sirtori E. (2016). *Forecasting the socio-economic impact of the Large Hadron Collider: A cost–benefit analysis to 2025 and beyond*. Technological Forecasting and Social Change 112: 38-53.

Giacomelli, P., Bianchin, C., Iconomidou-Fayard, L., Sciascia, B., Niedziela, J. (2017) *CERN Impact on People's Careers*, presented at CERN Council Open Session on December 15th, 2017. https://indico.cern.ch/event/681590/contributions/2792974/attachments/15.

Giffoni, F., & Vignetti, S. (2019). *Assessing the Socio-economic Impact of Research Infrastructures: A Systematic Review of Existing Approaches and the Role of Cost-Benefit Analysis*. L'industria, 40(1), 75-102.

IOP (2012) The Career Paths of Physics Graduates: A Longitudinal Study 2006–2010. London: Institute of Physics (IOP).

Keeley, B. (2007). *Human capital, how what you know shapes your life*. OECD. https://www.oecd.org/insights/humancapitalhowwhatyouknowshapesyourlife.htm.

Jolliffe, I.T. and Cadima, J., (2016). Principal component analysis: a review and recent developments. Philosophical *Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065),. https://doi.org/10.1098/rsta.2015.0202

Lam, A. (2007). Knowledge networks and careers: Academic scientists in industry–university links. *Journal of management studies*, *44*(6), 993-1016.

Laurila, S. (2013). *Students as a bridge building link. In Student perspective on the cooperation of universities and large scale research facilities*. Presentation at NORDTEC-Conference (Lund, June)

Lee, S., & Bozeman, B. (2005). *The impact of research collaboration on scientific productivity. Social studies of science*, 35(5), 673-702.

Long J S and Freese J 2014 Regression Models for Categorical Dependent Variables Using Stata 3rd edn (College Station, TX: StataCorp LP).

Malone, T. W., & Bernstein, M. S. (Eds.). (2015). *Handbook of collective intelligence*. MIT Press.

Matsunaga, Y. (2015). Salary commensurate with social capital. *Applied Economics Letters*, *22*(14), 1116-1121.

Mincer, Jacob. (1974). *Schooling, Experience, and Earnings*. Human Behavior & Social Institutions No. 2.

Montenegro, C., Patrinos H. (2014). *Comparable estimates of returns to schooling around the world*. World Bank Group. Policy Research Working Paper 7020.

Mulvey, P. and Pold J. (2015). *Physics Doctorates Initial Employment Data from the degree recipient follow-up survey for the classes of 2013 and 2014*. American Institute of Phisics, 2016. www.aip.org/statistics.

Mulvey, P. and Pold J. (2017). *Physics Bachelors: Initial Employment*. American Institute of Phisics. https://www.aip.org/sites/default/files/statistics/employment/bachinitemp-p-14.1.pdf.

Nicholson., Mulvey P. (2014). *Trends in Physics PhDs*. American Institute of Phisics 2014, https://www.aps.org/careers/statistics/upload/trends-phd0214.pdf.

OECD (2014). *Report on the Impacts of Large Research Infrastructures on Economic Innovation and on Society*. Case studies at CERN. Paris. Global Science Forum report, OECD Publishing.

OECD (2019). *Reference framework for assessing the scientific and socio-economic impact of research infrastructures*. OECD Science, Technology and Industry Policy Papers, No. 65. https://doi.org/10.1787/3ffee43b-en

Peskin M., Prebys E., Ross M., Rappoccio S., Seiden A., Stroynowski R. (2013). *Benefits to the US from Physicists Working at Accelerators Overseas*. Minneapolis arXiv:1312.4884 : Electronic Proceeding Community Summer Study.

Pilbeam, C., Lloyd-Jones, G., & Denyer, D. (2013). Leveraging value in doctoral student networks through social capital. *Studies in Higher Education*, *38*(10), 1472-1489.

Pold, J., & Mulvey, P. (2015). *Physics Masters One Year after Degree: Results from the Follow-Up Survey of Master's Recipients*, Classes of 2012, 2013, & 2014 Combined. Focus On. AIP Statistical Research Center.

Psacharopoulos, G., & Patrinos, H. A. (2004). *Human capital and rates of return*. International handbook on the economics of education, 1-57.

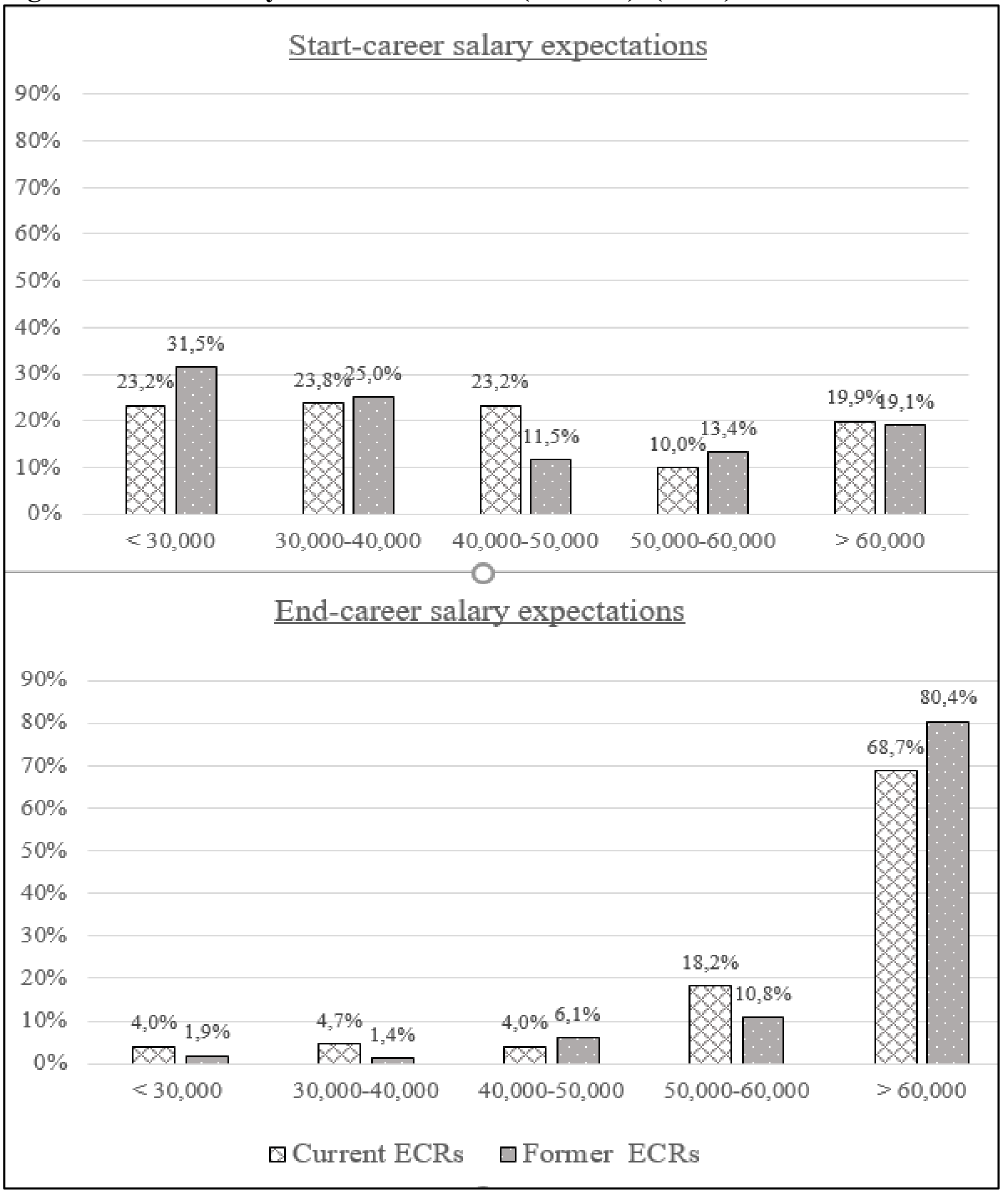
Scrivens, K., and C. Smith. (2013). *Four interpretations of social capital: an agenda for measurement*, OECD Statistics Working Papers, 2013/06

Schweitzer L, Lyons S, KJ Kuron L and SW Ng E (2014) *The gender gap in pre-career salary expectations: a test of five explanations* Career Dev. Int. 19 404–25

Tavakol, M. and R. Dennick (2011), 'Making sense of Cronbach's alpha,' *International Journal of Medical Education*, 2, 53–55.

Wolff, H. G., & Moser, K. (2008). *Choice, accountability, and effortful processing in escalation situations*. Zeitschrift für Psychologie/journal of Psychology, 216(4), 235.

Figure 1: Gross salary level distributions (N= 367[a]). (EUR)



Note: The statistical difference between the distribution of (expected) salaries between students and employees was assessed by using a Pearson's chi-square test for starting-career salary distribution and a Fisher's exact test for end-career salary distribution. The Fisher's test allows taking into account the low numbers of respondents (less than 5) in some categories of the end-salary expectations. Both the Pearson's chi-square test and The Fisher's test reject the null of the similarity of distributions. [a]71 missing answers to these questions.

Table 1 Type of skills acquired during the stay at CERN (N= 391[a]). Scale[b]: 1 =Not decisive; 5 = Highly decisive

Question B.7	Mean	% of respondents with "Decisive" or "Highly Decisive" responses
To what extent the following skills have been improved thanks to the experience at CERN LHC?		
Technical skills	4.28	86%
Scientific skills	4.24	85%
Problem-solving capacity	4.09	77%
Communication skills	3.96	72%
Independent thinking, creativity	3.95	72%
Developing, maintaining and using networks of collaborations	3.60	56%
Team/project leadership	3.53	52%

Note: [a]47 missing answers to this question. [b]Original Scale: 1 = Not decisive; 2 =Poorly decisive; 3=Moderately decisive; 4=Decisive; 5 = Highly decisive

Table 2 Main variables and definitions

Dependent Variable	**Definition**
End-career salary	It is coded as 1 if the elicited salary level is lower than EUR 30,000; 2 if it is between EUR 30,000 and 40,000; 3 if it is between EUR 40,000 and 50,000; 4 if it is between EUR 40,000 and 50,000; and 5 if the salary level is more than EUR 60,000;
Independent Variables	
Scientific human capital	It is an index which is positively linked to the two items in Table 1: "having improved technical skills" and "having improved scientific skills" thanks to the experience at CERN LHC. The index was obtained by performing a principal component analysis (PCA) as explained in detail in the Appendix,We expect that this variable positively correlates with salary expectations.
Social capital	It is a dummy variable, taking the value of 1 if the respondents chose the options "decisive" or "highly decisive" and 0 otherwise when answering the question in Table 1 "Developing, maintaining and using networks of collaborations."
Other soft skills	It is an index (principal component) which is positively linked to the remaining four items in Table 1: "communication skills", "developing leadership skills", "independent thinking and creativity", and "problem-solving capacity". It was calculated by the PCA as well (see Appendix).
Length of stay	It is a continuous variable which measures the duration in month of the experiential learning at CERN;
Male	It is a dummy variable, taking the value of 1 if the respondent was male and 0 otherwise;
Age	It is a continuous variable measured in years indicating the age of the respondent
PhD	It is a dummy variable taking the value of 1 if the respondent held a PhD or higher level of education and 0 otherwise;
Research	It is a dummy variable taking the value of 1 if the expected sector of activity is research (either in or outside academia) and 0 otherwise
Industry Finance	It is a dummy variable taking the value of 1 if the expected sector of activity is industry, finance or ICT and 0 otherwise;
Salary for comparators	It is a categorical variable capturing to what extent (according to respondents' beliefs) their salaries will be higher than the ones of their peers. It takes the value of 1 if the answer was 0% (no difference exists); 2 if 'up to 10%'; 3 if 'from 11 to 30%', and 4 if the answer was 'more than 30%'. The term 'Salary for comparators' is from Schweitzer et al. (2014), and we expect a positive impact of this variable on the levels of salary stated by respondents.

Alice, CMS, LHCb, Atlas, and *Other experiments*	They are a set of dummy variables identifying the experiment where respondents worked during their stay at CERN. They take the value of 1 if the respondent ticked that experiment and 0 otherwise.
Being former ECR	It is a dummy variable which assumes the value of 1 for former ECRs and 0 for students (current ECRs); This variable takes on board differences in the distribution of the stated salaries by current and former ECRs (employees).
Nationality	It is s set of dummies identifying the respondents' country of origin as earnings strongly depend on national labour market conditions and educational systems (Bianchin et al., 2019).
Face to face	It is a dummy variable that identifies the interview modality: 1 if it was carried out in person and 0 if it was online.

Table 3 Ordered Logit Model regressions. The dependent variable is End-career salary

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Scientific human capital	0.115* (0.069)		0.124* (0.070)	0.140** (0.073)	0.086 (0.154)	0.156* (0.088)	0.171* (0.088)
Social capital (*networking*)		0.095 (0.277)	0.069 (0.301)	0.190 (0.303)	0.170 (0.295)	0.031 (0.333)	0.024 (0.322)
Other soft skills		0.138 (0.142)	0.121 (0.141)	0.110 (0.144)	0.060 (0.126)	-0.186 (0.162)	-0.104 (0.152)
Length of stay				0.008* (0.005)	-0.008 (0.011)	0.018*** (0.006)	0.012** (0.005)
Length of stay*Scientific human capital					0.004** (0.002)		
Being former ECRs	0.772*** (0.283)	0.787** (0.293)	0.814*** (0.291)	0.515 (0.338)	0.819** (0.298)	0.613 (0.406)	0.659 (0.405)
Male						1.049*** (0.317)	1.042*** (0.304)
Age						-0.045 (0.029)	-0.034 (0.028)
PhD						0.051 (0.375)	0.208 (0.358)
Salary for comparators						0.555*** (0.166)	
Research						0.134 (0.383)	0.232 (0.375)
Industry_Finance						1.192*** (0.362)	1.237*** (0.355)
Experiment fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nationality fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Face-to-face	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	345	345	345	320	320	320	320
McFadden R2	0.035	0.034	0.038	0.040	0.039	0.122	0.100
Count R2	0.754	0.756	0.756	0.752	0.752	0.749	0.754
Log Likelihood	-276.620	-275.037	-274.073	-269.239	-269.410	-252.831	-259.608
Likelihood ratio test	20.311	19.494	21.422	22.041	24.941	61.943	51.896
Parallel regression assumption (p-value)	0.352	0.338	0.265	0.163	0.160	0.122	0.139

The table shows the drivers of the likelihood of choosing one of the end-salary categories. Robust standard errors in parentheses. ***, **, * denote significance at the 1%, 5% 10% level, respectively. For the sake of simplicity and as follows, we only report the estimates of the alpha cut-points in the full model (Column 6), while we do not report the other one referring to the other specifications oin the other columns, In the model above there are four cut-points for each specification and apart from some exceptions in Columns 3 and 4, they are always statistically significant justifying the use of five categories of the level of salary expectations over combining some categories. The

four estimated alpha cut-points in the full model are: $\alpha_1 = 0.277$; $\alpha_2 = 1.499$; $\alpha_3 = 2.206$; $\alpha_4 = 2.981$.
To be noted that the number of salary categories has been reduced from 10 (as asked in the questionnaire) to five to implement the econometric analysis.

Table 4 Marginal effects of length of stay (one additional month) on End-career salary

End-career salary category	Marginal effects (%)	Std. Dev
< 30,000 EUR	-0.060*	0.032
30,000-40,000 EUR	-0.036*	0.021
40,000-50,000 EUR	-0.037*	0.020
50,000-60,000 EUR	0.121**	0.056
> 60,000 EUR	0.253***	0.103
***,**,* denote significance at the 1%, 5% 10% level respectively.		

Table 5 Attractiveness for doctoral students to spend a research period at CERN according to team leaders

	Not important	Fairly important	Slightly important	Important	Very important	Total
Deepening the knowledge and competences in the domain of interest	1%	7%	0%	32%	60%	100%
Develop new professional skills	1%	8%	2%	40%	50%	100%
World undisputed prestige of CERN	3%	24%	7%	30%	36%	100%
Possibility to work with world class scientists and engineers	1%	5%	1%	33%	60%	100%
Working in an international environment	0%	5%	2%	28%	64%	100%

Table 6 ECR skills improved thanks to the experience at CERN as from team leaders' perceptions

	Not at all	To a small extent	To a moderate extent	To a high extent	To a very great extent	Total
Scientific skills	5%	2%	22%	48%	23%	100%
Technical skills	3%	4%	23%	41%	29%	100%
Communication skills	3%	3%	18%	43%	32%	100%
Problem-solving capacity	4%	9%	31%	39%	17%	100%
Team/project leadership	6%	10%	27%	33%	25%	100%
Developing, maintaining and using networks of collaborations	2%	3%	11%	40%	45%	100%
Independent thinking/critical analysis/creativity	6%	9%	33%	35%	17%	100%

Table 7 Team leaders expectations on students' salary premium (ranging from 4% to 12%)

Options	Respondents (%)
The range sounds reasonable to me	**54%**
I would have expected a greater impact	31%
I would have expected a lower impact	2%
I have no opinion	1%
I do not know at all	12%

Table 8 Salary premium for a doctorate degree in Science, Technology, Engineering and Mathematics: an

overview from secondary data sources

Reference	Description of data	Year	Statistics	Field of Science	Salary for a person with doctoral degree	Salary for a person with Master degree	Total Salary Premium	Total Premium (%)	1 year Salary Premium	1 year Salary Premium (%)
Doctorate:(Malvey and Pold, 2016) American Physical Society Statistical Research Center **Master: (Pold and Malvey, 2015) American Physical Society Statistical Research Center**	Doctorate: US - Data are based on respondents holding potentially permanent positions in the private sector (158) and in universities and 4-year colleges (36) and on postdocs in government labs (65) and in universities and UARIs (291) Master: US-figure based on the responses of 210 non-US citizens and 536 US citizens	Doctorate: 2013 & 2014 Master: 2012 & 2013 & 2014	Median (USD)	Physics	66,000 (government, 14%) 48,000 (university, 52%) 99,000 (private sector, 31%) Total: 71,000	41,000 (university, 53%) 65,000 (private sector, 19%) Total: 53,000	18,000	34%	4,500	8.5%
NACE (National Association of Colleges and Employers) Salary Survey	US - -NACE members (243 respondents)	2017	Average (USD)	Computer Science	110,841	81,039	29,802	37%	7,451	9.2%
				Engineering	95,973	75,053	20,920	28%	5,230	7.0%
				Math & Science	86,713	70,061	16,652	24%	4,163	5.9%
https://www.academics.com/science/salaries_who_earns_what_in_research_and_development_53193.html https://www.academics.com/science/what_chemists_earn_37951.html	Salary comparison conducted by Personal Market based on an evaluation of 15,857 datasets	2013	Average (EUR)	General scientific subject	55,300	48,000	7,300	15%	1,825	3.8%
				Natural Science	55,266	47,200	8,066	17%	2,017	4.3%
				Chemistry	63,000	57,000	6,000	11%	1,500	2.6%