European Journal of Physics

PAPER

Experiential learning in high energy physics: a survey of students at the LHC

To cite this article: Tiziano Camporesi *et al* 2017 *Eur. J. Phys.* **38** 025703

IOP Publishing | European Journal of Physics

Eur. J. Phys. **38** (2017) 025703 (22pp) doi:10.1088/1361-6404/aa5121

Experiential learning in high energy physics: a survey of students at the LHC

Tiziano Camporesi[1], Gelsomina Catalano[2], Massimo Florio[2,4] and Francesco Giffoni[3]

[1] CERN European Organization for Nuclear Research, Switzerland
[2] Department of Economics, Management and Quantitative Methods, University of Milan, Italy
[3] Centre for Economic and Social Research 'Rossi Doria', Roma Tre University, Rome, Italy

E-mail: massimo.florio@unimi.it

Received 20 November 2016
Accepted for publication 1 December 2016
Published 9 January 2017

CrossMark

Abstract

More than 36 000 students and post-docs will be involved until 2025 in research at the Large Hadron Collider (LHC) mainly through international collaborations. To what extent they value the skills acquired? Do students expect that their learning experience will have an impact on their professional future? By drawing from earlier literature on experiential learning, we have designed a survey of current and former students at LHC. To quantitatively measure the students' perceptions, we compare the salary expectations of current students with the assessment of those now employed in different jobs. Survey data are analysed by ordered logistic regression models, which allow multivariate statistical analyses with limited dependent variables. Results suggest that experiential learning at LHC positively correlates with both current and former students' salary expectations. Those already employed clearly confirm the expectations of current students. At least two not mutually exclusive explanations underlie the results. First, the training at LHC is perceived to provide students valuable skills, which in turn affect the salary expectations; secondly, the LHC research experience *per se* may act as signal in the labour market. Respondents put a price tag on their learning experience, a 'LHC salary premium' ranging from 5% to 12% compared with what they would have expected for their career without such an experience at CERN.

Keywords: experiential learning, salary expectations, human capital, Large Hadron Collider, CERN, early career researchers in physics

[4] Author to whom any correspondence should be addressed.

(Some figures may appear in colour only in the online journal)

1. Introduction

Students in physics are often involved in experiential learning in laboratories (Choi *et al* 2011). This is an effective way for them to gain practical knowledge and enhance their employment opportunities (Islam *et al* 2015). Employability may be further improved when such an experience takes place in a highly renowned laboratory, such as the European Organisation for Nuclear Research (CERN), (Schopper 2009).

CERN offers—mainly through the academic institutes associated to the experiments operating on the accelerator complex—several training opportunities. The programmes address respectively bachelor and master students, and PhD students in physics, engineering and computing. Typically, the undergraduate programmes give the possibility to spend at CERN 4–14 months; while in the doctoral programmes 6–36 months. Post-docs at experiments are then offered for several years. The number of incoming PhD students and post-docs at LHC experiments was about 9000 from 2009 (first LHC run) to 2014. Florio *et al* (2016) forecast this number to be 36 800 in the 1993–2025 period (17 400 students and 19 400 post-docs)[5].

By analysing the career of physics students involved at the Delphi experiment at CERN's Large Electron-Positron Collider from 1982 to 1999, Camporesi (2001) suggested that the interest of the private sector in students and researchers who spent a period at CERN 'cannot be in the knowledge of fundamental law of nature, but rather on the skills that our students acquire. [...] Whatever they do go on to do, their stay at CERN certainly plays a major role' (p 146). In the same vein, OECD (2014) emphasises that 'the intellectual environment at high-energy physics (HEP) laboratories is exceptional, and is probably comparable to that of the most innovative high-technology companies' (p 18). Students in such an environment improve their skills by working on experiments, interacting with different cultures, writing their PhD thesis, participating to meetings, conference and workshops. These competencies can be exploited in many workplaces, even outside HEP (Camporesi 2001, Boisot 2011, OECD 2014). Yet, through a survey targeted to the US High-Energy Physics community, Anderson *et al* (2013) confirm that many of the skills learned in the laboratories are valued much both on academic and non-academic career path (see also Danielsson 2013, Laurila 2013).

More in general, according to earlier literature on salary expectations, graduates have own expectations about their professional lives based on their information set (Shelley 1994). Van Maanen and Schein (1977) define careers as a sequence of experiences and transitions. As a result, expectations individuals form before entering in the labour market or at the entry level, do influence their decisions about the next steps of their whole professional life. This hypothesis has been empirically validated by demonstrating that perceptions and the information set at pre-career level strongly affect subsequent salary increases (Keaveny and Inderrieden 2000, Fernandez-Mateo 2009). In line with this literature, the participation in HEP experiments may increase the human capital of students and positively influence their

[5] The taxonomy adopted by CERN classifies students into the following categories: doctoral students (mostly from institutes participating to CERN based Experiments or directly supported by CERN for specific Applied Physics programmes), CERN technical students, CERN fellows and Users. See CERN Personnel Statistics yearly reports for details. The figures reported here only refer to the apportionment of these personnel categories to the LHC. Users and Fellows aged more than 35 as well as participants to summer schools or short courses are not included (see Florio *et al* 2016 for details).

professional expectations. Training at international collaborations may improve technical and problem-solving capacity as well as team-work capabilities, management and communications skills. The latter have been often found poor in science graduates without such advanced experimental training (Rodrigues *et al* 2007, Sharma *et al* 2007, O'Byrne *et al* 2008, Institute of Physics (IOP) 2012, Nielsen 2014).

While earlier research suggests that salary expectations are influenced by experiential learning in international collaborations, there is not yet a coherent explanation on why and how this relationship between training and reward expectations arises. Focussing on the LHC, this paper attempts to fill this gap by answering a set of research questions: To what extent the experiential learning at LHC is valued and affects expectations of students, after controlling for the personal characteristics and other potential confounding factors? Which of the acquired skills mediate the relationship between this experience and professional expectations? And to what extent? How much a perceived 'LHC premium' is worth? Are the perceptions of current students aligned with those of former students who are now employed in different fields?

In order to answer these questions, we interviewed 384 students and former students at LHC. Interviews were collected by means of a questionnaire-based survey carried out between October 2014 and March 2015 through face-to-face interviews at CERN and on-line questionnaire. Then, we performed a multivariate statistical analysis of the data.

The remainder of this paper is structured as follows. Section 2 describes the research methodology. Specifically, this section introduces a conceptual model linking the experience as student or post-doc at the LHC and their expectations. Section 3 presents the results, considering starting and end-career salary expectations, both of current and former students. Section 4 concludes.

2. Research methodology

Having in mind our research questions, we developed a survey of both current and former students at LHC based on a structured questionnaire (Camporesi *et al* 2016). Current students are respondents who, at the time of the interview, were involved in different international collaborations at the LHC, particularly at CMS. Hereafter, we refer to them simply as students. In contrast, former students are those individuals who, after having been students at the LHC, at the time of the survey either worked at CERN[6] or they had left CERN and were employed in different jobs, including outside science. Hereafter, we refer to them as former students or equivalently as employees.

The questionnaire was structured along four sections. The first two sections inquired about personal information and experience at LHC. They were targeted to both students and former students. Section three focused on students and it investigated on expectations about their professional career including starting and end-career salary expectations. The fourth section was directed to former students only and inquired about both the current professional career and future expectations. Clearly, the starting salary of former students refers to their

[6] E.g. as users, fellows, or associates. 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. Fellows are graduates of a higher educational establishment, typically with a maximum of ten years' relevant professional experience. They are appointed by the CERN for a limited period of time to perform functions within the CERN as part of their professional development. Cooperation Associates are scientists, technicians and engineers admitted by CERN to contribute on behalf of their home institution to the execution of a collaboration under an agreement between the CERN and their home institutions (see CERN Personnel Statistics yearly reports for details, or visit http://useroffice.web.cern.ch).

Table 1. PCA and MCA results and descriptions of commonalities between questions for each factor.

Original question Col. 1	Items Col. 2	Loadings (<0.3 left blank) Col. 3	Factor score (*label used in the order logistic regressions*) Col. 4	Eigenvalue Col. 5	% of variance explained Col. 6
Considering that your time at LHC is equal to 100%, please indicate the % dedicated to the following activities:	Participation to meetings/dealing with coordination activities (e.g. managing working groups, etc)	0.45	Factor 1	2.24	37
	Participation to conferences and workshops	0.37			
	Participation to other training activities	0.31			
	Outreach activities (e.g. guide to visitors)				
	Working on experiments (e.g. data analysis)	0.37	Factor 2	1.15	19
	Writing thesis/papers/articles	0.82			
	(KMO = 0.59)		**% of cumulated variance = 56**		
How do you rate the importance of the following considerations on your decision of applying for a research period at LHC?	World undisputed prestige of CERN	0.60	Factor 1 (*Networking motivation*)	2.11	42
	Possibility to work with world class physicists	0.52			
	Working in an international environment	0.58			
	Deepening the knowledge and competences in the scientific domain of interest	0.79	Factor 2 (*Skill motivation*)	1.07	21
	Develop new professional skills	0.59			

Table 1. (Continued.)

Original question Col. 1	Items Col. 2	Loadings (<0.3 left blank) Col. 3	Factor score (*label used in the order logistic regressions*) Col. 4	Eigenvalue Col. 5	% of variance explained Col. 6
	(KMO = 0.68)		**% of cumulated variance = 63**		
To what extent the following skills have been improved thanks to the experience at LHC?	Scientific skills	0.55	Factor 1 (*Technical skills*)	2.51	36
	Technical skills	0.55			
	Problem-solving capacity	0.44			
	Independent thinking/critical analysis/creativity	0.43			
	Communication skills	0.44	Factor 2 (*Communication skills and leadership*)	1.98	28
	Developing, maintaining and using networks of collaborations	0.60			
	Team/project leadership	0.64			
	(KMO = 0.84)		**% of cumulated variance = 64**		
Please indicate the expected sector of your future career	Industry	0.50	Factor 1 (*Future Sector*)	1.44	65
	ICT sector (e.g. computing)	0.45			
	Financial sector	0.54			
	Public administration				
	Research (at CERN and other than CERN)	−0.31			
	University and other teaching	−0.42			
	(KMO = 0.57)		**% of cumulated variance = 65**		
Please indicate the expected position of your future career	Manager		Factor 1	1.61	42
	Engineer	0.76			
	Data Analyst	0.63			

Table 1. (Continued.)

Original question Col. 1	Items Col. 2	Loadings (<0.3 left blank) Col. 3	Factor score (*label used in the order logistic regressions*) Col. 4	Eigenvalue Col. 5	% of variance explained Col. 6
	Physicist	0.65	Factor 2	1.14	23
	Professor/Researcher	0.58			
	(KMO = 0.53)		**% of cumulated variance = 65**		

first or current professional experience and thus, it is an observed salary and not an expectation.

Except for salary expectations, questions related to future outlooks utilise multiple-item constructs, measured with two different types of scale: ordinal and nominal. Ordinal scales employ five-point Likert scales, with anchors of 1 and 5, indicating the weighting assigned by individuals to a set of not mutually exclusive statements about their working experience at LHC. Nominal-type scales differentiate between multiple items based on qualitative classifications such names or meta-categories. Nominal variables were coded as binary (1/0) variables.

In order to homogenise the available information without loss of relevant statistic information as well as obtain new continuous variables (factor scores) to constitute the inputs for later multivariate analysis, two techniques were used for statistical pre-treatment of data. The first one was factor analysis of principal components (hereafter, PCA) applied to those questions measured by ordinal Likert scales; the second one was factor analysis of multiple correspondence (hereafter, MCA) applied to questions having nominal items[7].

Table 1 shows the results of this exercise. Columns 1 and 2 report the original questions and their items, respectively. Column 3 shows loadings which indicate the correlations between each factor (Column 4) and the observable items[8]. Factor scores are reported in Column 4. They are the main output of the PCA and are indices that combine the information in the items. Our factors' labelling is reported in brackets in Column 4[9]. Finally, Columns 5 and 6 show the eigenvalue associated with each factor and the percentage of items' variance explained by each factor, respectively[10]. The Kaiser–Meyer–Olkin (KMO) value at the end of each question establishes that the extracted factors in the PCA account for most of the variance in responses.

We tested the influence of LHC experiential learning on both starting and end-career range of salary expectations (our limited dependent variables) by using ordered logistic regressions. Differently from linear regression model, the ordered logistic model is a non-linear regression model for ordinal limited dependent variables. Let y_i be our dependent variable measuring the range of salary expectations and taking integer values from 1 to J. Suppose, also, that the underlying process to be analysed is:

$$y_i^* = X_i\beta + \varepsilon_i,$$

where y_i^* is the exact but unobserved (latent) dependent variable (i.e. the exact level of agreement with the statement proposed by the interviewer), X_i is a vector of independent

[7] The suitability of the data for PCA was tested for each question by using the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy with the threshold value of 0.5 (Cheung and Yeung 1998, Cheung *et al* 2000). The number of factors to include in later multivariate analyses was determined according to Kaiser's (1961) rule of thumb suggesting the retention of those factors with an eigenvalue greater than unity. In addition, Hair *et al* (1998) suggest that, in social science, factors may be stopped at least when 60 per cent of the cumulative variance was explained. Loadings were detected to interpret the principal component solution. As for the MCA, we made use of the Greenacre's (1993) formula to select the relevant factors (see also Abdi and Valentin, 2007). The interpretation of factors was based on their graphical projections (Blasius and Greenacre 1998, Greenacre 2000). For further details see an earlier version of this paper available at https://arxiv.org/abs/1607.01941. For another recent application of PCA to a survey addressed to physics students see Mason and Singh (2016).

[8] Loadings lower than 0.3, indicating weak correlation, were not reported in table 1.

[9] Some of factors identified at this stage were not retained in the later multivariate analysis because they were found not statistically significant in explaining salary expectations. Thus, no labels were assigned to them.

[10] If the eigenvalue drops below 1, it means that the factor explains less variance than a single item; namely, it does not provide any additional information than that contained in the single item. Thus, only the factors that better explain the items' variance are retained.

variables (see below) aiming at explaining the range of salary expectations, β is the vector of regression coefficients we wish to estimate and ε_i is the random disturbance term that follows a logistic distribution (Balakrishnan 1992). The variable y_i relates to the latent variable (y_i^*) according to the rule:

$$y_i = 1 \quad \text{if} \quad y_i^* \leqslant \tau_1,$$

$$y_i = j \quad \text{if}\ \tau_{j-1} < y_i^* \leqslant \tau_j \quad j = 2, \ldots, J-1,$$

$$y_i = J \quad \text{if}\ \tau_{J-1} < y_i^* < \infty,$$

where $\tau_1 \leqslant \tau_2 \leqslant \cdots \leqslant \tau_{J-1}$ are unknown thresholds (cut-points) to be estimated. The conditional distribution of y_i given X_i is given by:

$$\Pr(y_i = j | X_i) = \Lambda(\tau_j - X_i\beta) - \Lambda(\tau_{j-1} - X_i\beta),$$

where Λ (.) denote the logistic cumulative distribution function. The above equation tells us what is the probability that the respondent selects one of the proposed range of salary expectations given the value of the independent variables. Empirically, this is investigated by estimating the marginal effects of such variables on this probability. The beta-coefficients are estimated by using maximum likelihood procedure (see Long and Freese 2014 for further details).

Drawing on contemporary research on salary expectations (Maihaus 2014, Schweitzer *et al* 2014, Frick and Maihaus 2016) and on science (mainly, physics) graduates job market (Sharma *et al* 2008, Hazari *et al* 2010, Jusoh *et al* 2011, IOP-Institute of Physics 2012, Nielsen 2014, Islam *et al* 2015), we identified the following four sets of independent variables.

Set 1. *Personal characteristics.*

- *Male*. It is a dummy variable taking on the value 1 for males and 0 for females.
- *Age* is a continuous variable measured in years.
- *PhD* is a dummy variable taking on the value 1 if the highest education qualification is at least a PhD or the PhD is on-going; and 0 for master and bachelor degrees.
- *Nationality* is a dummy variable which takes on value 1 if the respondent comes from a CERN Member State and 0 otherwise[11].
- *Physics* is a dummy variable, which takes on value 1 if the academic background is physics and 0 otherwise (e.g. engineering or computer science).
- *Employee* is a dummy variable, which takes on value 1 if the respondent is an employee and 0 if he is a student.

Set 2. *Experience at LHC.*

Respondents were asked to what extent the following skills have improved thanks to the experiential learning at LHC:

- *Technical skills*. It is a continuous variable (factor score, see table 1) which is linked to skills such problem-solving capacity, scientific and technical skills, independent thinking, critical analysis and creativity.

[11] CERN Member States at the time of the survey were Austria, Belgium, Bulgaria, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Italy, Netherlands, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, United Kingdom, and Israel.

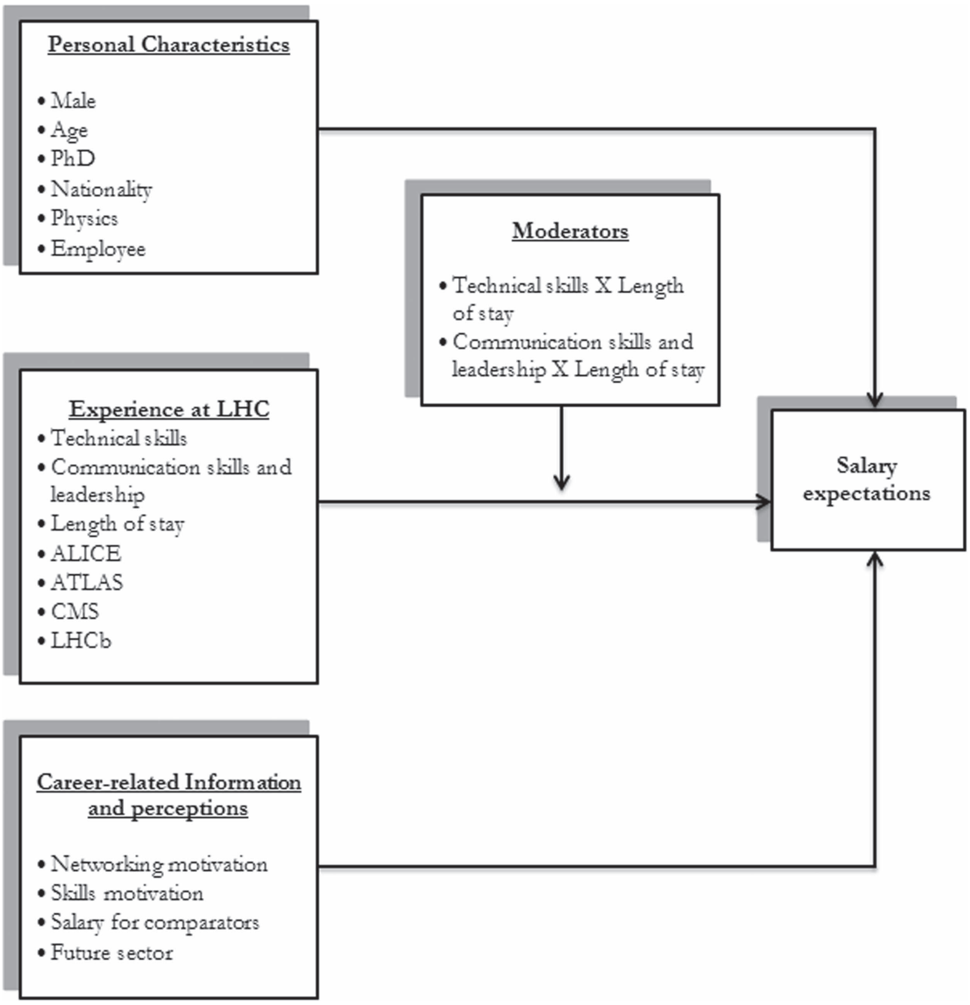

Figure 1. Conceptual model.

- *Communication skills and leadership*. It a continuous variable (factor score, see table 1) and it is related to skills such as communication, team/project leadership, developing, maintaining and using networks of collaborations.

The length of the research period spent at LHC and the type of experiments respondents have worked on are also proxies of the experiential learning at LHC. As a result, we consider:

- *Length of stay*. It indicates the length of the research period individuals have spent at LHC. It is a continuous variable measured in months.
- *ALICE, ATLAS, CMS, LHCb*. They identify the four main experiments of the LHC. Each of the experiments is codified as a dummy variable taking on the value of 1 if the respondent has worked on that experiment and 0 otherwise.

  It can be argued that the longer is the stay at LHC, the more likely that students develop valuable skills, which in turn increase their pay expectations. Using this assumption, we introduce the following:

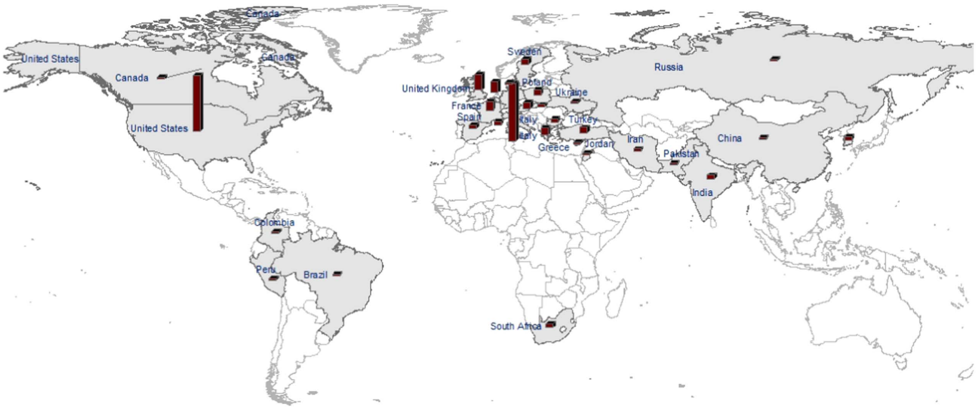

Figure 2. Share of respondents by nationality.

Set 3. *Moderators.*

- *Technical skills X Length of stay*. It is an interaction term between the length of the research period individuals have spent at LHC and technical skills. It is a continuous variable.
- *Communication skills and leadership X Length of stay*. It is an interaction term between the length of the research period individuals have spent at LHC and communication skills.

Set 4. *Career-related information and perceptions*.

- *Networking motivation*. It is a continuous variable (factor score, see table 1) and it is related to the importance of *networking* in the decision of applying for a research period at LHC. The greater the value, the more important was for respondents to apply because of the possibility to work with world-class physicists and in a prestigious and international institution as CERN is.
- *Skill motivation*. Unlike the previous variable, this factor score (see table 1) is linked to the relevance of developing personal and professional skills rather than to networking.
- *Salary for comparators*[12]. It is a categorical variable, which describes to what extent respondents expect that their future salary will be higher than that earned by their peers. It takes on value 1 if 0%, 2 if up to 10%, 3 if 11%–30%, 4 if more than 30%.
- *Future sector* is a continuous variable (factor score) which is positively linked to sectors such industry, finance, and ICT and negatively related to research and university.

Summing up, we introduce a comprehensive model to test the relationships between the experience at LHC—proxied by the skills acquired and/or length of stay—(Camporesi 2001, Boisot 2011, OECD 2014, Florio *et al* 2016) and range of salary expectations (limited dependent variables), by controlling for personal characteristics (Set 1), career-related information and perceptions (Set 4) and the type of experiments, which individuals have worked on (i.e. ALICE, ATLAS, CMS, LHCb). Furthermore, we test the hypothesis according to which the predictive effect of the skill acquired at LHC and the length of stay may interact each other (Set 3) meaning that the longer is the stay at LHC, the more likely that students develop valuable skills, which in turn increase pay expectations. Our final point is to identify the value that students attach to such a working experience. To this end, we look at

[12] We borrowed this terminology from Schweitzer *et al* (2014).

Table 2. Personal characteristics and length of stay at LHC.

Variable	Total ($n = 318$)	Students ($n = 141$)	Employees ($n = 177$)
Discrete Variables			
Gender (%)			
Male	73.3	70.9	75.1
Female	26.7	29.1	24.9
Education (%)			
At least PhD	71.4	48.2	89.9
Less than PhD	28.6	51.8	10.1
Nationality (%)			
Member State	62.3	61.7	63.3
Non-Member State	37.4	38.3	36.7
Academic background (%)			
Physics	85.5	80.9	89.3
Other	14.5	19.1	10.7
Continuous Variables			
Age (years)			
Mean	31.1	28.2	33.4
Std. Dev.	4.7	3.5	4.0
Min	21	21	25
Max	44	38	44
Length of stay at LHC (months)			
Mean	44.7	24.4	59.3
Std. Dev.	34.7	17.4	36.6
Min	1	1	1
Max	181	72	181

the marginal effects of the experiential learning spent at LHC on salary expectations. The model is shown in figure 1.

3. Data and results

The survey was carried out between October 2014 and March 2015 and addressed to current and former students working on LHC experiments. It resulted in 384 questionnaire collected, of which 221 through face-to-face interviews at CERN and 163 filled in online[13]. Respondents come from 52 countries, mainly from Italy (22%), USA (16%) followed by Germany (8%), UK (7%), France, Belgium and Greece (4%, each) (figure 2). About 63% of the sample is from a CERN Member State. Because of missing data in some interviews, the final sample used for our analysis includes 318 valid questionnaires (195 collected face-to-face and 123 online).

Table 2 reports some descriptive statistics. Males represent 73% of the sample (71% amongst students and 75% amongst former students, here simply labelled as employees). 71% of respondents have at least a PhD as their highest education level; the remainder are bachelor

[13] Additional details of the survey are available in Catalano *et al* (2015).

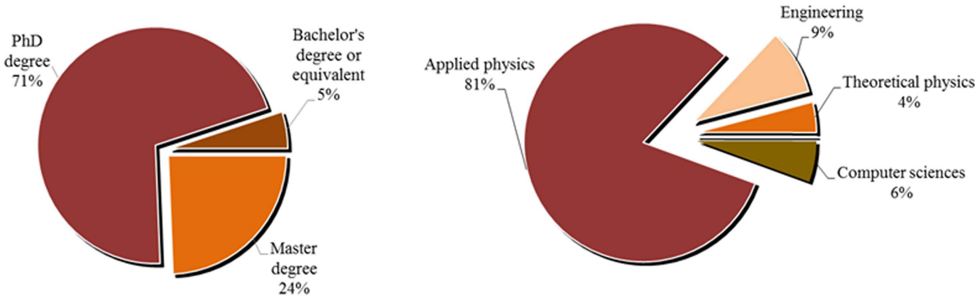

Figure 3. Share of respondents by educational degree and academic background.

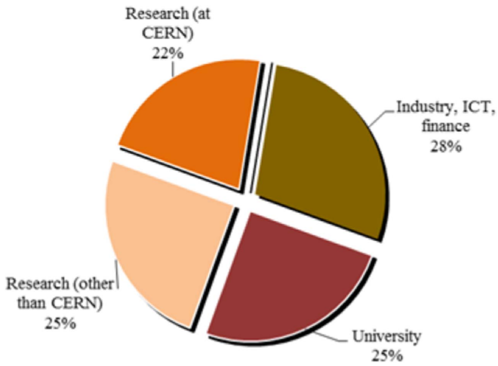

Figure 4. Employment sector. Share of employees.

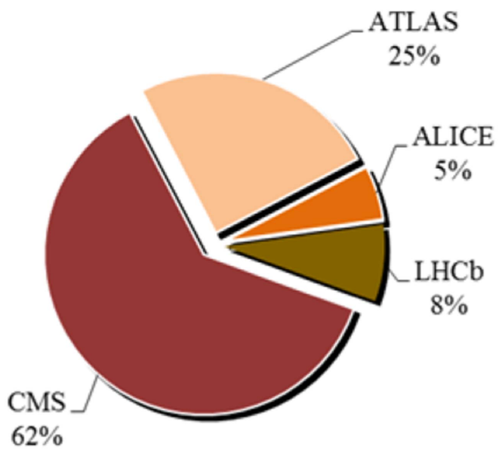

Figure 5. Respondents by LHC experiments.

or master degree holders. Amongst the employees, the percentage of those with a PhD is 90%. Actually, they are mostly post-docs.

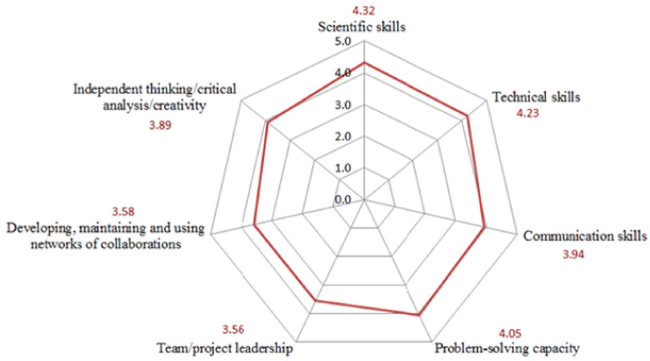

Figure 6. Skills improved thanks to the LHC.

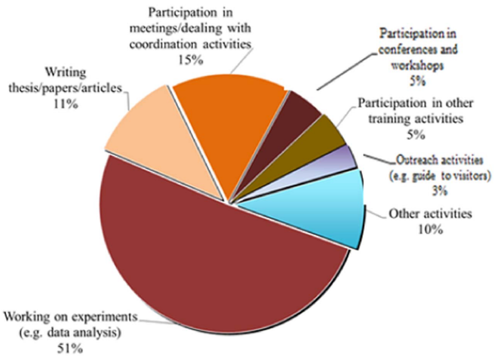

Figure 7. Time spent distribution across activities.

With regard to the academic background, 85% are physicists, while the remaining 15% have a degree in engineering or computer sciences (figure 3).

The average age of respondents is 31 years with an average of 45 months working experience at LHC. Among students (average age equal to 28 years) the average training period is 24 months, while among former students (average age equal to 33 years), the average length of stay at LHC is about 60 months. Finally, the distribution of respondents among the different sectors and experiments is shown in figures 4 and 5 respectively. Note that in our sample, CMS is over-represented because the survey was first launched with the CMS collaboration. Afterwards, the survey was extended to other experiments.

Figures 6 and 7 report, respectively, descriptive statistics of the variable related to skills acquired at LHC and the kind of activity on which respondents have spent most of the time during such an experience.

Figure 6 displays that, according to respondents, the LHC experience has improved their technical skills more than communication and leadership skills, while figure 7 shows that most of the time respondents have spent at LHC, was dedicated to working on experiments, and specifically, data analysis (51%) and writing papers and/or thesis (11%).

The (unconditional) distribution of starting and end-career gross salary expectations split by employment status is reported in table 3.

Table 3 shows that responses about starting career expected salaries tend to group in the lowest salaries categories (less than EUR 50 000) for current students more than for former

Table 3. Gross salary expectations distributions (percentage).

	Starting-career salary		End-career salary	
Category	Students (%)	Employees (%)	Students (%)	Employees (%)
<30 000 EUR	22.7	13.1	4.3	2.3
30 000–40 000 EUR	24.1	23.2	5.1	1.7
40 000–50 000 EUR	24.1	14.9	3.6	5.2
50 000–60 000 EUR	10.1	16.7	19.6	11.6
>60 000 EUR	19.1	32.1	67.4	79.2

Note. Employment status differences in the expected salaries were 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. Both tests reject the null of the similarity of distributions.

students (who observe the actual level). The distribution of end-career salary expectations is concentrated in the highest categories (more than EUR 50 000) for both students and employees, with again the latter more optimistic. In order to test whether students and employees differ in their expected salaries, we carried out a Pearson's chi-square test in the initial salary case and a Fisher's exact test in the case of end career salaries expectations[14]. The chi-square test ($p < 0.01$) and the Fisher's exact test ($p < 0.05$) suggest that there is a statistically significant difference between students and employees in expected salaries. We control for such dissimilarity in the following multivariate analysis by including an employment status dummy variable. The dissimilarity in this preliminary analysis suggests that salary expectations are to a certain extent higher for former students than for current students. We assume that the employees have gained more actual information on pay in the job market, at least for the entry level, than students.

Table 4 analyses the overall correlation between the variables entering in the conceptual model[15]. Statistically significant correlations are observed between salary expectations (both starting and end-career expectations) and *Technical Skills* and *Length of stay* respectively; in contrast, *Communication skills/leaderships* do not correlate with salary expectations. Personal characteristics such as *Male* and *Employees* and career-related perceptions such as *Salary for comparators* and *Future sector* positively correlate with salary expectations as well.

To answer our research questions, in principle, we may use both starting salary or end-career salary expectations as dependent variables (Schweitzer *et al* 2014). Actually, in our sample, they are strongly and positively correlated ($\text{coef} = 0.62,\ p < 0.05;$ table 3) suggesting that using two different regression analysis would not lead us to notably different conclusions. In addition, European Commission (2014, chapter 7) suggests that the benefit of human capital development should be measured on the lifelong salary. Therefore, we only make use of end-career salary expectations as dependent variable[16]. Results are shown in

[14] The use of two different tests is necessary because the chi-square test assumes that the value of each cell is five or higher. While this assumption is met in the distribution of starting salary expectations, it does not hold for expected salary at peak of careers.
[15] The variables associated to the type of experiment ALICE, ATLAS, CMS and LHCb do not show any significant correlation with the relevant variables we are interested in (i.e. salary expectations, technical skills, communication skills/leadership and length of stay). Thus, they are not reported in the table.
[16] Regressions, which make use of starting-career salary expectations as dependent variable are available upon request.

Table 4. Correlations matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Starting career salary	1														
2. End career salary	0.62**	1													
Personal characteristics															
3. Male	0.23**	0.19**	1												
4. Age	0.19**	0.02	0.11**	1											
5. PhD	0.08	0.04	0.01	0.33**	1										
6. Nationality	−0.08	−0.07	−0.08	−0.03	−0.11**	1									
7. Physics	−0.09	−0.14**	−0.05	−0.20**	0.40**	−0.17**	1								
8. Employee	0.18**	0.12**	0.04	0.56**	0.46**	0.02	0.12**	1							
Experience at LHC															
9. Technical skills	0.10**	0.12**	−0.07	0.00	0.07	−0.03	0.00	0.02	1						
10. Communication skill/leadership	−0.01	−0.05	−0.01	0.20**	0.10	0.00	−0.04	0.15**	−0.01	1					
11. Length of stay	0.20**	0.14**	0.01	0.40**	0.33**	0.14**	0.12**	0.50**	0.12**	0.05**	1				
Career-related information and Perceptions															
12. Networking motivation	−0.03	0.02	−0.11**	0.01	−0.08	0.16**	−0.09	0.00	0.00	0.38**	−0.07	1			
13. Skill motivation	0.02	−0.00	−0.08	−0.04	0.01	0.02	0.10	−0.14**	0.37**	−0.20**	−0.10**	0.29**	1		
14. Salary for comparators	0.11**	0.15**	−0.01	−0.07	−0.03	−0.04	−0.12**	−0.03	0.16**	−0.03	−0.02	0.18	0.10	1	
15. Future sector	0.15**	0.18**	0.10	−0.01	−0.01	0.01	−0.09	0.12	−0.00	−0.13**	0.12	−0.16**	− 0.19**	0.03	1

Note. The variables associated to the type of experiment ALICE, ATLAS, CMS and LHCb do not show any significant correlation with the relevant variables we are interested in (i.e. salary expectations, technical skills, communication skills/leadership and length of stay). Thus, they are not reported in the table. ** Significant at 5% level.

Table 5. Ordered logistic estimates. Dependent variable is end career salary expectation.

Variables	(1) coef	se	(2) coef	se	(3) coef	se	(4) coef	se	(5) coef	se
Experience at LHC										
Technical skills	0.103^{*}	(0.062)			0.110^{*}	(0.061)	0.004	(0.145)	0.135	(0.134)
Length of stay			0.009^{*}	(0.005)	0.009^{*}	(0.005)	0.011^{**}	(0.005)	0.017^{**}	(0.007)
Technical skills X Length of stay							0.004^{**}	(0.002)	0.004^{**}	(0.002)
Personal Characteristics										
Employee	0.814^{***}	(0.282)	0.455	(0.346)	0.493	(0.352)	0.500	(0.354)	0.444	(0.409)
Male									0.946^{***}	(0.349)
Age									−0.035	(0.043)
PhD									2.653^{***}	(0.924)
Physics									−0.294	(0.449)
Career-related information										
Networking motivation									−0.098	(0.157)
Skill motivation									0.272	(0.239)
Salary for comparators									0.342^{***}	(0.130)
Future sector									0.495^{***}	(0.155)
Nationality-specific effects	Yes		Yes		Yes		Yes		Yes	
Experiments-specific effects	Yes		Yes		Yes		Yes		Yes	
Interview-specific effects	Yes		Yes		Yes		Yes		Yes	
Observations	318		318		318		318		318	
McFadden's R2	0.036		0.035		0.043		0.050		0.159	
Log Likelihood	−254.3		−240.8		−237.4		−235.9		−172.8	
Likelihood ratio test	16.87		17.99		19.17		22.75		52.20	
Proportional odds hp test (*p*-value)	0.291		0.276		0.227		0.205		0.182	

Table shows the determinants of the probability of falling in one of the expected salary category. Robust standard errors in parentheses. ***, **, * denote significance at the 1%, 5% 1% level, respectively.

table 5. For each of the regression proposed, the proportional odds assumption, underlying ordered logistic procedure, was tested (Long and Freese 2014, chapter 7)[17]. The *p*-values are reported in the last row of the table.

We carried out the analysis in five steps. In the first step (Column 1) we only include the types of skills respondents declared having improved thanks to the training at LHC, which is one of our proxy of the experiential learning. Actually, in the regressions, we only included the variable (factor score) *Technical skills*; the variable *Communication skills and leadership* was never found statistically significant. In the second step (Column 2), we test the length of the research period as proxy of training at LHC; in doing so, we exclude *Technical skills* and include *Length of stay*. The third step (Column 3) shows that *Technical skills* and *Length of stay* remain significantly associated with salary expectations also when both variables are jointly plugged into the same model. As mentioned, one may argue that skills acquired at HEP experiments increase or improve as the length of the research period increases. We test this hypothesis in the fourth step (Column 4) by adding the *moderators*. The fifth step (Column 5) presents the full model, which controls for personal characteristics, career-related information and perception of respondents. We failed to find any statistical evidence on the contribution of additional interaction terms on salary expectations[18].

Regardless the step, we always control for four types of specific-effects: first, the employment status (employee versus student). It enables us to capture unobserved heterogeneity that may shape salary expectations of such individuals beyond the experience at LHC. Second, we consider nationality-fixed effects. To the extent that individuals form their salary expectations according to some features of the country of origin, for example labour market conditions or the prevailing type of educational system (Hazari *et al* 2010, Wickramasinghe and Perera 2010, Jusoh *et al* 2011, Maihaus 2014), this dummy should capture such an effect. Third, we consider experiments-specific effects. These dummies identify the experiments at which respondents have spent their training period at LHC: ALICE, ATLAS, CMS and LHCb[19]. Last but not least, we include interview-specific effect[20]. It allows us to reduce any systematic difference between responses obtained by personal interviews and through online questionnaire (Duffy *et al* 2005).

Column 1 and Column 2 reveal that experience at LHC positively and significantly correlates with salary expectations both when it is proxied by the acquired competences and by *Length of stay*. These variables keep their statistically significance up also when they are plugged simultaneously into the same model (Column 3), suggesting that the time spent at LHC generates *per se* increasing salary expectations, aside the skills acquired.

[17] One of the assumption underlying order logistic regression is that the relationship between each pair of outcome group is the same. Put differently, the proportional odds assumption requires that the coefficients describing the relationship between, let's say, the lowest versus all higher categories of the response variable are the same as those that describe the relationship between the next lowest category and all higher categories, etc. Because the relationship between all pairs of groups is the same, there is only one set of coefficients (only one model); otherwise, a generalised ordered logistic model should be run. In order to test the proportional odds assumption, we run the Brant test, rather than the 'omodel' command, since the latter does not recognised categorical variables. The null hypothesis is that there is no difference in the coefficients between models. In our case, the proportional odds assumption is met in all of proposed regressions, except for the first model (Column 1, table 5). For further details see Long and Freese (2014) or visit http://ats.ucla.edu/stat/stata/dae/ologit.htm.

[18] In an unreported regression, several interaction terms between personal characteristics and career-related information and our proxies of experience at LHC (Technical skills and Length of stay) were tested as suggested by Hogue *et al* (2010). We found no statistical significance. However, results are available upon request.

[19] Even though these dummies variables could be potential interesting for the purpose of our analysis, we found them never statistically significant.

[20] This dummy variable takes on the value of 1 if the interview was carried out face-to face and 0 otherwise.

Column 4, adds the interaction term between *Technical Skills* and *Length of stay*. The positive and statistically coefficient (coef $= 0.004$, $p < 0.05$) indicates that the skills acquired at LHC increases as the time spent on the experiments increases, which in turn generates higher rewards expectations. This is confirmed by the fact that the variable *Technical Skills* loses its predictive power in explaining salary expectations. As before, *Length of stay* retains its own significance (coef $= 0.011$, $p < 0.01$).

The estimated association between salary expectations and experiential learning at LHC remains robust also after adding respondents' personal characteristics as well as their career-related information and perceptions (Column, 5). In addition, the coefficient on *Male* is positive and statistically significant at 1 per cent level reflecting a substantial gender gap in salary expectations among graduates (Ng and Wiesner 2007, Hogue *et al* 2010, Schweitzer *et al* 2014), and, particularly among physicists (Hazari *et al* 2010, Lissoni *et al* 2011, IOP-Institute of Physics 2012). The variable PhD enters positively and significantly as well, confirming that salary expectations increase with educational attainment (Shelley 1994, Jusoh *et al* 2011, Islam *et al* 2015). Interestingly, in this multivariate context there are no more significant differences on end-career salary expectations between employees and students; this result means that once controlling for personal characteristics, the employment status loses its predictive power in explaining end-career expected salaries. This is probably due to the fact that, after all, the community of HEP is relatively small and information circulates amongst researchers of different seniority, at least for not too distant cohorts.

Column 5 also shows that the variables *Salary for comparators* and *Future sector* enter into the model with a significant and positive coefficient. The *Salary for comparators* variable suggests that the higher the salary respondents are expected to earn with respect to their comparable peers thanks to their research experience at LHC, the higher their own salary expectations are (Schweitzer *et al* 2014). As regard *Future sector,* higher salaries are expected in sectors such as industry and finance; in contrast, respondents expect lower salaries in academia. Finally, the likelihood ratio tests in the models indicate that the variation in the independent variables explains a good proportion of the variability in the response variable[21].

In order to assess the 'LHC premium', we look at marginal effects of the working experience at LHC (proxied by the *Length of stay*) on end-career salary expectations. If a premium is expected, then it should be measured on end-career salary expectations (European Commission 2014 chapter 7, Schweitzer *et al* 2014, Florio *et al* 2016). Marginal effects are those stemming from the full model (Column 5, table 4) and they are shown in table 6, where values are reported in percentage terms.

One additional month of training spent at LHC increases the probability of declaring an expected salary in the two highest categories (50 000–60 000 EUR and >60 000 EUR) and reduces the probability of expecting a low salary (less than 50 000 EUR). For example, an additional month of experiential learning at LHC increases the probability of expecting a salary greater than 60 000 EUR by 0.27 percentage points, ceteris paribus.

Let us now focus on the two highest categories (respectively marginal effect 0.125% and 0.268%, $p < 0.05$; $p < 0.01$), which contain almost 85% of responses. Note also that, in our sample, the average number of months spent at LHC is 44 for the whole sample, 24 for students and 60 for employees. Thus, for an 'average' individual who declared an expected

[21] For the sake of simplicity, we chose to not include the constants (i.e. the taos parameters) of the regressions in table 5. In the ordered logistic models, the constants (here, we have four constants for each model) are cut-points used to differentiate the adjacent levels of the dependent variable. Apart from some exceptions in Columns 2 and 3, they were found all statistically significant, justifying the use of five categories of the level of salary expectations over combining some categories. Actually, some preliminary elaboration on original data leads us to reduce the salary expectation categories from ten to five.

Table 6. Marginal effects of *Length to stay* (one additional month) on End-career salary expectations.

End-career salary expectations categories	Marginal effects (%) coef	se
<30 000 EUR	−0.066*	(0.035)
30 000–40 000 EUR	−0.038*	(0.021)
40 000–50 000 EUR	−0.039*	(0.021)
50 000–60 000 EUR	0.125**	(0.057)
>60 000 EUR	0.268***	(0.108)

***,**,* denote significance at the 1%, 5% 10% level respectively.

salary between EUR 50 000 and 60 000 the experiential learning at LHC is worth about 5% excess salary (3% for a student and 7% for an employee). For those respondents whose expected salary falls in the category '> 60 000', the stay at LHC is worth, on average, about 12% (6% for a student and 16% for an employee). This is the range of our final estimation of the expected 'LHC premium' as perceived by current and former students, based on their information set, after controlling for personal characteristics, country of origin, etc.

4. Conclusions

This paper contributes to the literature on experiential learning by a statistical analysis of perceptions of students in a HEP laboratory. Survey data were collected from students and former students at the largest particle accelerator in the world, the Large Hadron Collider. We were particularly interested in understanding to what extent earlier results by Camporesi (2001) on students involved in experiments at LEP, a previous major collider at CERN, are confirmed for more recent cohorts of students. Moreover, and this is the main novelty of our contribution, we wanted to measure quantitatively the intensity of perceptions about the learning experience at the LHC by estimating an expected salary premium. Then, we wanted to study the drivers of such expectations.

There are several reasons why this context is of interest for a broader research perspective on professional expectations in science careers. The LHC operates at the frontiers of science, and for this reason it attracts students from a very large number of countries (over 50 in our sample). This fact ensures that possible specific country effects play a minor role. The research community of particle physics can be considered as a relatively small but dense global social network, where information on career opportunities is widely shared within each cohort and across cohorts of early career researchers. Moreover, there is fragmentary but interesting evidence that students at CERN will have a professional future in a variety of jobs, beyond academic research, including in industry and finance. Thus, it seems that the LHC context, including its experiments (such as CMS that hosts many European but also US students) is an ideal testing ground for the more general question of the experiential learning in physics.

There are three main findings of our analysis. First, after controlling for possible confounding factors, there is no statistical difference in end-career expectations between the two sub-samples of respondents: current and former students. In fact the latter, who have acquired more direct information, are slightly on average more optimistic in their perceptions of the salary premium, but the difference is not statistically significant after controlling for individual characteristics. This suggests that the research community actually shares the information on professional opportunities and this fact shapes relatively homogeneous expectations.

This is also indirect evidence of realism of the expectations, because for former students they are based on actual job market information. Moreover, one may argue that if respondents have greater self-confidence and skills upon exiting CERN, then they are likely to demand and receive higher salaries[22]. A second finding is that the core drivers of the expectations are *length of stay* at the LHC and *technical skills* acquired. Hence, the perceived professional premium is not attributed to a purely reputational effect in the job market associated with the mere fact of having been selected for training at CERN, but it increases proportionally to the time spent in research in that context. Respondents were able to indicate on a five-point scale which were the most important skills acquired: the salary premium increases according to the perceived importance of technical skills. This result clearly points to the perception of experiential learning as a driver of professional opportunities. Finally, the interaction between the two drivers is statistically significant as well.

Third, we can conclude that, according to the convergent perception of respondents, either current or former students at CERN (the latter now employed in a variety of occupations, including industry and finance), there is a professional premium arising from sustained experiential learning in the laboratory. The premium is estimated in the range 5%–12% over the entire career, compared to peers not having had the opportunity to being involved in the LHC experiments. In other words, the laboratory environment is at the same time perceived by insiders as a scientific discovery machine and as an engine of human capital formation, i.e. a 'career springboard' as initially suggested by Camporesi (2001). For the first time we have now been able to measure this expectation from within.

Acknowledgments

The authors are very grateful to all the interviewees for their participation in the survey and to the CMS Collaboration for the initial input in its development. An earlier version of this paper has been produced in the frame of the project 'Cost/Benefit Analysis in the Research, Development and Innovation Sector' sponsored by the EIB University Research Sponsorship programme (EIBURS), whose financial support is gratefully acknowledged (http://eiburs.unimi.it/). Additional research in 2016 has been carried out in collaboration with CERN in the context of the Future Circular Collider Study (http://cern.ch/fcc) that develops concepts for a European post-LHC energy-frontier accelerator research infrastructure. This paper should not be reported as representing the official views of the EIB or the CERN, or any third party. We are grateful for comments on an earlier version to Johannes Gutleber (FCC Study), Stefano Forte (University of Milan), Giovanni Zevi della Porta (University of California, San Diego) and to Stefano Carrazza (CERN). Any errors remain those of the authors.

References

Abdi H and Valentin D 2007 Multiple correspondence analysis *Encyclopedia of Measurement and Statistics* ed N J Salkind (Thousand Oaks, CA: Sage) pp 651–7

Anderson J 2013 Snowmass 2013 Young Physicists Science and Career Survey Report arXiv:1307.8080

Balakrishnan N 1992 *Handbook of the Logistic Distribution* (New York: Marcel Dekker)

Blasius J and Greenacre M 1998 *Visualization of Categorical Data* (New York: Academic)

Boisot M 2011 *Collisions and Collaboration: The Organization of Learning in the ATLAS Experiment at the LHC* (Oxford : Oxford University Press)

[22] We are grateful to an anonymous reviewer for this comment.

Camporesi T 2001 High-energy physics as a career springboard *Eur. J. Phys.* **22** 159–148

Camporesi T, Catalano G, Florio M and Giffoni F 2016 A 'LHC premium' for early career researchers? perceptions from within arXiv:1607.01941

Catalano G, Del B C and Florio M 2015 Human capital formation at LHC: survey results (http://wp.demm.unimi.it/tl_files/wp/2015/DEMM-2015_10wp)

Cheung S O and Yeung Y W 1998 The effectiveness of the dispute resolution advisor system: a critical appraisal *Int. J. Project Manage.* **16** 367–74

Cheung S O, Tam C M, Ndekugri I and Harris F C 2000 Factors affecting clients' project dispute resolution satisfaction in Hong Kong *Constr. Manage. Econ.* **18** 281–94

Choi S H J, Funk M, Roelofs S H, Alvarez-Elizondo M B and Nieminen T A 2011 International research work experience of young females in physics arXiv:1101.1758

Danielsson H 2013 Students as a bridge building link *CERN Presentation at NORDTEC Conference (Lund, June)* https://lth.se/fileadmin/lth/images/Nyhetsbilder/NORDTEK_Hans_Danielsson.pdf

Duffy B, Smith K, Terhanian G and Bremer J 2005 Comparing data from online and face-to-face surveys *Int. J. Market Res.* **47** 615–39

European Commission 2014 Guide to cost-benefit analysis of investment projects *Economic Appraisal Tool for Cohesion Policy 2014–2020* (Brussels: European Commission)

Fernandez-Mateo I 2009 Cumulative gender disadvantage in contract employment1 *Am. J. Sociol.* **114** 871–923

Florio M, Forte S and Sirtori E 2016 Forecasting the social impact of the Large Hadron Collider: a cost-benefit analysis to 2025 and beyond *Technol. Forecast. Social Change* **112** 38–53

Frick B and Maihaus M 2016 The structure and determinants of expected and actual starting salaries of higher education students in Germany: identical or different? *Educ. Econ.* **24** 374–92

Greenacre M 2000 Correspondence analysis of square asymmetric matrices *Appl. Stat.* **49** 297–310

Greenacre M J 1993 *Correspondence Analysis in Practice* (London: Academic)

Hair J F Jr, Anderson R E, Tatham R L and Black W C 1998 *Multivariate Data Analysis* 5th edn (Englewood Cliffs, NJ: Prentice-Hall)

Hazari Z, Sonnert G, Sadler P M and Shanahan M C 2010 Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: a gender study *J. Res. Sci. Teach.* **47** 978–1003

Hogue M, DuBois C L and Fox-Cardamone L 2010 Gender differences in pay expectations: the roles of job intention and self-view *Psychol. Women Q.* **34** 215–27

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

Islam T, Ahmed I, Khalifah Z, Sadiq M and Faheem M A 2015 Graduates' expectation gap: the role of employers and higher learning institutes *J. Appl. Res. Higher Educ.* **7** 372–84

Jusoh M, Simun M and Choy Chong S 2011 Expectation gaps, job satisfaction, and organizational commitment of fresh graduates: roles of graduates, higher learning institutions and employers *Educ. Train.* **53** 515–30

Kaiser H F 1961 A note on Guttman's lower bound for the number of common factors *Br. J. Stat. Psychol.* **14** 1–2

Keaveny T J and Inderrieden E J 2000 Gender differences in pay satisfaction and pay expectations *J. Manage. Issues* **12** 363–79

Laurila S 2013 Students as a bridge building link. Student perspective on the cooperation of universities and large scale research facilities *Presentation at NORDTEC-Conference (Lund, June)* https://lth.se/fileadmin/lth/images/Nyhetsbilder/NORDTEK_Santeri_Laurila.pd

Lissoni F, Mairesse J, Montobbio F and Pezzoni M 2011 Scientific productivity and academic promotion: a study on French and Italian physicists *Ind. Corp. Change* **20** 253–94

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

Maihaus M 2014 *The Economics of Higher Education in Germany: Salary Expectations, Signaling, and Social Mobility* (Marburg: Tectum Verlag)

Mason A J and Singh C 2016 Surveying college introductory physics students' attitudes and approaches to problem solving *Eur. J. Phys.* **37** 055704

Ng E S and Wiesner W H 2007 Are men always picked over women? The effects of employment equity directives on selection decisions *J. Bus. Ethics* **76** 177–87

Nielsen T L B 2014 From master's programme to labour market: a study on physics graduates' of the transition to the labour market *Master Thesis* University of Copenhagen

O'Byrne J, Mendez A, Sharma M, Kirkup L and Scott D 2008 Physics graduates in the workforce: does physics education help? *Congress of the Australian Institute of Physics* (Australian Institute of Physics)

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

Rodrigues S, Tytler R, Darby L, Hubber P, Symington D and Edwards J 2007 The usefulness of a science degree: the 'lost voices' of science trained professionals *Int. J. Sci. Educ.* **29** 1411–33

Schopper H 2009 *LEP—The Lord of the Collider Rings at CERN 1980-2000: The Making, Operation and Legacy of the World's Largest Scientific Instrument* (Berlin: Springer)

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

Sharma M, Pollard J, Mendez A, Mills D, O'Byrne J, Scott D and Low D 2007 What does a physics undergraduate education give you? A perspective from Australian physics *Eur. J. Phys.* **29** 59

Shelley K J 1994 More job openings—even more new entrants: the outlook for college graduates, 1992–2005 *Occup. Outlook Q.* **38** 5–9

Van Maanen J E and Schein E H 1977 *Toward a Theory of Organizational Socialization* (Cambridge, MA: Sloan School of Management)

Wickramasinghe V and Perera L 2010 Graduates', university lecturers' and employers' perception towards employability skills *Educ. Train.* **52** 226–44