APPLIED ECONOMICS

Editor MARK TAYLOR
Co-Editor DAVID A. PEEL

APPLIED ECONOMICS LETTERS

Editor MARK TAYLOR
Co-Editor DAVID A. PEEL

Projecting the socio-economic impact of a Big Science Center: the world's largest particle accelerator at CERN

Andrea Bastianin[*]

(andrea.bastianin@unimi.it)

Department of Economics, Management and
Quantitative Methods, Università degli Studi di Milano

Chiara F. Del Bo

(chiara.delbo@unimi.it)

Department of Economics, Management and
Quantitative Methods, Università degli Studi di Milano

Massimo Florio

(massimo.florio@unimi.it)

Department of Economics, Management and
Quantitative Methods, Università degli Studi di Milano

Francesco Giffoni

(giffoni@csilmilano.com)

CSIL, Centre for Industrial Studies.

This draft: October, 12 2022

Abstract. Public investment in Big Science generates social benefits that can ultimately support economic growth. This paper implements a model for the Social Cost–Benefit Analysis (CBA) of Big Science and relies on Monte Carlo methods to quantify the uncertainty of long-term projections. We evaluate social costs and benefits of the High Luminosity upgrade of the Large Hadron Collider (HL-LHC) at CERN up to 2038. Monte Carlo simulations show that there is a 94% chance to observe a positive net present value for society. The attractiveness of CERN for Early Stage Researchers and technological spillovers for collaborating are key for a positive CBA result. Cultural effects, especially those related to onsite visitors, also contribute to generating societal benefits.

JEL classification: C15; D61; J24; O31; O33.

Keywords: cost-benefit analysis; Monte-Carlo methods; CERN; Big Science; Technology transfer

---

[*] *Corresponding author*: Department of Economics, Management and Quantitative Methods, University of Milan. Via Conservatorio, 7. I-20122 - Milan – Italy.
Email: andrea.bastianin@unimi.it

1. Introduction

Investing in science yields a return for entities that directly fund and manage R&D activities, while also generating additional social benefits (Del Bo, 2016; Griliches, 1958). A positive differential between the private and social rate of return justifies policies that support R&D investments as engines of economic growth (see e.g. Griliches, 1992; Jones and Williams, 1998; Nelson and Romer, 1996). Big Science Centres (BSC) producing research in high energy physics, astronomy, molecular biology, material and medical sciences yield societal benefits for different stakeholders.[1] These include scientists, Early Stage Researchers (ESR), firms involved in high-tech procurement, users of open software and data, consumers of downstream innovative products and citizens taking advantage of "cultural goods" (see e.g. Autio et al., 2004, Bastianin et al. 2021, Battistoni et al., 2016; Castelnovo et al., 2018; Del Bo et al. 2016; Florio et al. 2016; OECD, 2014; Scaringella and Chanaron, 2016; Vuola and Hameri, 2006).

An important methodological question is how to quantify the risk surrounding long-run projections of costs and benefits of BSC and large-scale research infrastructures[2] (RI), such as particle colliders. While researchers can prove the existence of a scientific motivation for upgrading an existing RI or building a new one, investors – Member States or participating institutions - want an assessment of the riskiness of the project over different phases, from construction to operation. A key risk is associated, for instance, with cost overruns. These risks are well known in the social Cost Benefit Analysis (CBA) of major infrastructure projects (see e.g. Johansson and Kristrom 2018), but have yet to be studied in the context of Big Science. The history of the demise by the US Congress of the Superconducting Super Collider in 1991 after a huge cost overrun shows that the construction of scenarios for Big Science can be subject to significant mistakes (Riordan et al., 2015), and that a

---

[1] "Big Science" identifies the style of scientific research characterizing fields such as high energy physics, astronomy and molecular biology after World War II (Dennis, 2017). BSCs typically rely on large-scale research instruments and infrastructures, have several participating institutions, attract funding from governments and use public procurement to develop technologies required for scientific research (see e.g. Linton, 2008; Florio, 2019).

[2] The European Strategy Forum on Research Infrastructures defines RI as "*facilities, resources or services of a unique nature identified (...) to conduct top-level research activities in all fields*." [ESFRI, 2016, p. 1]

sound methodology to account for and address the underlying risk is needed. This paper addresses the question of how to quantitatively model the risk surrounding costs and societal benefits when evaluating large RI with a CBA.

We contribute to the literature on the socioeconomic impacts of investing in Big Science by illustrating a methodological strategy rigorously applied to in-depth case study. First, we show how to implement Monte Carlo methods as part of a CBA of Big Science projects. Our approach delivers a relatively simple, yet powerful, method to quantify the risks of such projects, despite the uncertainty surrounding their scientific results. Our approach solves the puzzle of quantitatively assessing risk in a CBA of any major project that has an intrinsically uncertain - hence non-quantifiable - outcome. Our strategy is straightforward: we disregard Knightian uncertain outcomes related to discoveries, which, by definition, cannot be modelled (Knight 1921) and focus on risks that can be modelled with probability distributions, thus providing a test of whether the expected social benefits are greater than costs, even leaving the non-quantifiable unknown aside. Second, we apply our methodology to a case study, the socioeconomic impact of the new major project of the European physics community: the High Luminosity upgrade of the Large Hadron Collider (HL-LHC), the world's largest and most powerful particle accelerator. We stress that our approach can be applied to the CBA of any major Big Science project, from space exploration to biomedical research, provided that the analyst carefully selects and calibrates a set of key parameters pertaining to the stochastic component of the model. Importantly, our approach can be applied both to new infrastructures and to upgrades of existing ones.

Application to projecting the socioeconomic impacts of the HL-LHC project suggests that there is a 94% chance to observe a positive net present value of socio-economic benefits with an associated Monte Carlo error of 0.012. This fact points to a clear, quantifiable economic benefit for society of this investment. This is an interesting result *per se*, but more in general, it is an example of how long-run projections of Big Science projects should be assessed.

Our paper contributes to the literature on the design of long-term projections and forecasts that are crucial for policy design, portfolio allocation, decision-making and planning. (e.g. Congressional Budget Office, 2007 for an application to US economic aggregates; Lee 2011 and Raferty and Ševčíková, 2021 for population projections; Colby and Len, 2016 and Florio et al., 2018a for a focus on particle accelerators). The literature has developed a wide range of tools to produce long-term forecasts and asses their uncertainty that rely on both on time series methods (Granger and Jeon, 2007; Müller and Watson, 2017) and judgmental approaches (Lawrence et al., 2006; Önkal et al., 2013). Judgmental and scenario-based forecasting are particularly useful in settings such as the one considered in this paper, where lack of historical data renders the use of statistical methods unfeasible. Our approach is thus similar to the one pursued in population projections where a variety of methods, including reliance on experts with domain knowledge, are combined to produce projections over very long-run horizon. Such projections - that extend far into the future - cannot be directly evaluated relying on standard measures of forecasts performance, however their uncertainty can be assessed using Monte Carlo methods as we do in this paper and as it is done in other contexts (see e.g. Long and Mcmillen, 1987; Pflaumer, 1988).[3]

It also relates to the literature on the economic effects of BSC (see, e.g., Autio et al., 2004; Helmers and Overman, 2017; Vuola and Hameri, 2006). Lastly, it relates to a growing strand of literature on the socioeconomic impacts of CERN, a benchmark case study for the analysis of BSC (see, e.g., Åberg, and Bengtson, 2015; Bianchi-Streit et al.,1984; Bastianin and Del Bo, 2020; Bastianin et al., 2021; Castelnovo et al., 2018; Florio et al., 2018b; Nilsen and Anelli, 2016; OECD, 2014). An important contribution in this literature is Florio et al. (2016) who perform an ex-post CBA analysis of the LHC. Our work differs from Florio et al. (2016) in two major respects. First, our study extends Florio et al. (2016) because it deals with the hi-luminosity upgrade of the LHC, an already existing RI. Focusing on an upgrade, rather than on a new RI, poses an important methodological challenge,

---

[3] From now on we use the term projections for our results to distinguish them from forecasts based on historical data.

that is the development of a complex and detailed counterfactual scenario absent in Florio et al. (2016). Second, we show how to exploit Monte Carlo methods to produce policy-relevant conditional scenarios that constrain the value of selected key variables and hence can be used to as robustness check of the main conclusions. In particular, given the likelihood of cost overruns in BSC, we develop a pessimistic cost scenario showing that even in this case the probability of benefits exceeding costs remains substantially high.

Section 2 introduces CERN and the HL-LHC project. Section 3 describes a CBA model for the evaluation of RI, while Sections 4 and 5 provide implementation details and illustrate results. Section 6 concludes.[4] An Appendix with further details on data and methods completes the paper.

2. The world's most powerful particle collider

CERN is at the French-Swiss border and is built around a complex of accelerators, *i.e.* a succession of machines that accelerate particles to increasingly higher energies. It represents an international RI whose mission is to push the frontiers of science and technology with the goal of advancing knowledge in fundamental physics (Nielsen and Anelli, 2016). As of 2020, CERN's budget amounts to about 1,200 MCHF per year and over 17,000 people world-wide are involved (CERN, 2020). CERN is run by 23 Member States that contribute to capital and operating costs of its programs, and express their votes to approve plans of activity, the budget and review expenditures.

The LHC is the last and largest element of the accelerator complex supplying it with protons via a chain of four smaller accelerators that boost the particle beams and divide them into bunches. The LHC is a 27-kilometer ring of superconducting magnets housed in an underground tunnel where two beams of particles, traveling at almost the speed of light, are made to collide at four locations hosting the detectors and CERN's main experimental collaborations. Experiments at the LHC use detectors to analyse particles produced by collisions in the accelerator. Each of these experiments - run by

---

[4] For further details: https://fcc-cdr.web.cern.ch and Bastianin and Florio (2019).

collaborations of international scientists - has a distinct scientific objective. The four biggest experiments, located in huge underground caverns along the LHC ring, are: ATLAS, CMS, ALICE and LHCb.[5] Each experiment uses specific detectors for focussing on different phenomena. Experimental activity at the LHC started in September 2008 and in 2012 led to the discovery of the Higgs boson (Evans,2016).

A key indicator of an accelerator's performance is "luminosity" that measures the numbers of potential collisions per unit of time. The amount of data gathered by experiments is proportional to an accelerator's luminosity, making it a proxy of its discovery potential. Since the phenomena that physicists are looking for are incredibly rare, higher luminosity is needed to increase the probability of observing such processes. Over the period 2008-2018, the LHC has been operated at increasing luminosity over subsequent experimental runs. Statistical gains from running further experiments without considerably increasing the luminosity are marginal[6] and hence an upgrade of the LHC project – known as High-Luminosity LHC, (HL-LHC or HILUMI) - was announced as the top priority of the European Strategy for Particle Physics since 2013.[7] The aim of the HL-LHC project is to boost the luminosity of the LHC by a factor of ten and to increase the number of collisions by a factor of between five and seven.[8] This upgrade is key for escalating the discovery potential of the LHC and to extend its lifetime while the European physics community agrees on what will be its goal for the post-LHC era (Brüning and Rossi, 2015).

The HL-LHC project began in 2011, while construction began in 2015 with the publication of a technical design report. The civil-engineering work started in April 2018 and is expected to end in

---

[5] The smallest experiments instead are: TOTEM, LHCf, MoEDAL and FASER.

[6] As an example of how luminosity affects the performance of an accelerator and hence its discovery potential, CERN reports that "*the running time necessary to half the statistical error in the measurements will be more than ten years at the end of 2019*" (https://hilumilhc.web.cern.ch/content/lhc-baseline).

[7] The European Strategy for Particle Physics identifies the open, inclusive and science-driven process that the European physics community adopts to plan and prioritize its long-term research agenda (https://europeanstrategy.cern/).

[8] To put these numbers in perspective, we note that CERN reports that the HL-LHC "*will produce at least 15 million Higgs bosons per year, compared to around three million from the LHC in 2017*" (https://home.cern/science/accelerators/high-luminosity-lhc).

2022. Installation of the first components began during the second "Long Shutdown" of the LHC,[9] in the 2019-2021 time period; the remaining equipment and major experiment upgrades will be installed during the third Long Shutdown, between 2025 and 2027. Experiments are then expected to run for at least 12 years, up to 2040. For the purpose of the model presented in Section 3, we refer to the period between 2008 and 2014 as the LHC baseline, while in 2015 the HL-LHC upgrade starts.

The HL-LHC project is led by CERN with the support of an international collaboration of 29 institutions in 13 countries. Upscaling luminosity involves several technological challenges, as well as civil-engineering work to adapt the existing LHC tunnel and infrastructures to the new design of the accelerator. New equipment must be installed over about 1.2 kilometres of the LHC tunnel and major technological advances are needed at the sites of two experiments: ATLAS and CMS. These include a set of more powerful magnets made of a superconducting compound (i.e. niobium-tin) that will be used for the first time in an accelerator with the aim of focusing the particle bunches before they meet. Moreover, innovative superconducting power lines - able to carry currents of record intensities - approximately one hundred metres long, will connect the power converters to the accelerator. Boosting the HL-LHC's performance also demands work along the chain of four upstream accelerators – including the development of a new linear accelerator, Linac4 - that pre-accelerates the beams before sending them into the largest ring. To install such new equipment substantial civil-engineering work is required. Shafts of around 80 metres will be dug on the sites of the ATLAS and CMS experiments where two service tunnels will also be built, along with five additional surface buildings.

Building the HL-LHC pushes several technologies – including superconductors, vacuum technologies, computing, electronics - and industrial processes to their limits. Some of these innovations will likely be translated into benefits beyond basic research. Superconducting magnets already find applications in the fields of medical imaging and cancer treatment with particle beams

---

[9] Experiments at the LHC are interrupted on a regular schedule to allow maintenance and upgrading works at the accelerator. Such periods are called "long shutdows".

(see e.g. Battiston et al. 2016). Other studies explore the possibility of using the HL-LHC superconducting power lines to transport, in an environmentally-sustainable way, high electrical power over great distances.[10] Last, but not least, these technological advances rely on the continuing training of new scientist at CERN, including physicists, engineers, data scientists and technicians. This amounts to creating new human capital that may be employed elsewhere, including outside physics.

The scientific outcome of the HL-LHC is intrinsically uncertain. Nobody knows if the larger number of Higgs bosons and other particles created in the collisions will advance knowledge beyond what physicists already know: the Standard Model of Physics (see, e.g., Gaillard et al., 1999), a powerful set of equations describing fundamental interactions of nature, but unable to explain e.g. gravity, prevalence of matter against anti-matter and other puzzles. Nevertheless, an attempt to quantitatively assess the expected socio-economic impacts of such an investment need to be made also to assist further developments of RI at CERN and elsewhere. In fact, CERN has recently unveiled the plan to build the Future Circular Collider (FCC), the replacement of the HL-LHC. This project entails the construction of a 100-kilometer long underground tunnel to house the new collider with construction costs in the range 9.5-21.8 billions Euro.[11] Construction and design of the FCC collider should take up to 23 years, with operation not starting before 2040 and experiments are expected to be run for 15-25 years.

3. A CBA framework for risk assessment of Big Science projects

The framework for the CBA of RI developed by Florio and Sirtori (2016) can be adapted to the study of societal costs and benefits of any Big Science project, and is thus our starting point.

---

[10] See: https://home.cern/resources/faqs/high-luminosity-lhc

[11] These figures are sourced from Abada et al. (2019) and converted from Swiss Francs to Euro using the CHF/EUR exchange rate of 0.91 recoded on 30 April 2020. The width of the cost range depends on the type of collider that will be housed in the new tunnel. We consider two scenarios the construction of the FCC-*ee* collider and the standalone implementation of the FCC-*hh* collider. See Bastianin and Florio (2019) for a non-technical summary of different FCC scenarios.

There are several reasons for focusing on the CBA methodology. First, CBA is deeply rooted in economic theory (Boardman et al., 2017; Florio, 2014; Johansson and Kriström, 2018) and well suited to quantitatively assess the multi-dimensional outputs of basic research carried out in BSC (Martin 1996). [12] Second, CBA is routinely applied to real-world evaluation problems (see e.g. Brent, 2017; Greenberg et al., 2013; Maresova et al., 2017). For instance, the "*Horizon 2020 – Work Programme 2018-2020 on European Research Infrastructures*"[13] mentions that the preparatory phase of new European Strategy Forum on Research Infrastructures (ESFRI) projects should include a CBA (European Commission, 2017). Major projects financed with the European Regional Development Fund and the Cohesion Fund[14] were explicitly required, up until the 2014-2020 programming period, to provide a full-fledged CBA following the methodology proposed in European Commission (2014). In the current programming period 2021-2027, while not a formal requirement, national Managing Authorities must provide information on the methods used to select projects eligible for funding (European Commission, 2021). The European Investment Bank uses CBA methods to decide the allocation of its funds among different projects, including RI.[15]

3.1 Methodology

The key output of a CBA model is the estimate of the expected Net Present Value (NPV) of the project over a given time horizon:[16] $E(NPV_j) \equiv E[DB_j - DC_j]$. The subscript identifies project j,[17] while DB_j and DC_j denote the cumulative sum of discounted social benefits and costs, respectively, of the RI over its lifespan. For new projects, for which a counterfactual (CF) scenario is not available - or alternatively, if the CF scenario corresponds to a "*do nothing*" case, that is, not building the RI -

---

[12] Alternative qualitative and quantitative methodologies are reviewed by Salter and Martin (2001) and Giffoni et al. (2021).

[13] https://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-infrastructures_en.pdf

[14] According to Article 100 of EU Regulation No 1303/2013, a major project is an investment operation comprising "*a series of works, activities or services intended to accomplish an indivisible task of a precise economic and technical nature which has clearly identified goals and for which the total eligible cost exceeds EUR 50 million*."

[15] https://www.eib.org/en/publications/economic-appraisal-of-investment-projects

[16] Two related measures of performance that are often presented are the Economic Rate of Return (i.e. the rate of return that yields zero NPV) and the "*B/C ratio*" (i.e. the ratio between discounted economic benefits and costs).

[17] In our case, j refers either to the *HL-LHC* project or to the counterfactual (*CF*) scenario.

interpretation of results is straightforward: a RI passes the CBA test when benefits exceed its costs for society, that is when the expected NPV is greater than zero. A CF scenario is instead necessary when comparing alternative projects for the same RI or when a "*do nothing*" option cannot be invoked. In these cases, the CBA test focuses on incremental costs and benefits for society and relies on the expected value of the incremental NPV: $\Delta NPV_{j,CF} = NPV_j - NPV_{CF}$. It follows that project j is preferred to the CF project if $E(\Delta NPV_{j,CF}) > 0$.

The practical implementation of the model presented above and the evaluation of risk involves three steps: (*i*) identification of the social costs and benefits associated with a given RI; (*ii*) their projection over the project horizon; (*iii*) the use of Monte Carlo methods to approximate the probability distribution functions of costs, benefits and hence of the *NPV* of the project.[18] With an estimate of the distribution of the *NPV* we can compute several statistics that can inform decision makers, such as the probability of observing a positive *NPV*.

Each RI affects different stakeholders and yields a distinct set of social benefits and costs. Within the CBA framework, costs include operating and investment expenditures (DF_j) as well as indirect costs for the society (DX_j). Any emission of pollutants, noise, traffic jams related with construction and operation of the RI are examples of negative externalities. As for benefits, we identify the following:

- value of training, or human capital formation, for students and early-stage researchers (DH_j);
- technological spillovers for suppliers, collaborating firms and other economic agents (DT_j);
- cultural effects for the general public (DL_j);
- value of academic publications and pre-prints for scientists (DS_j);
- existence or public good value of the RI (DE_j).

---

[18] From now on, for ease of notation we drop the subscript "*j*" referring to a particular scenario.

A more precise definition of each category of costs and benefits for the HL-LHC is provided below.

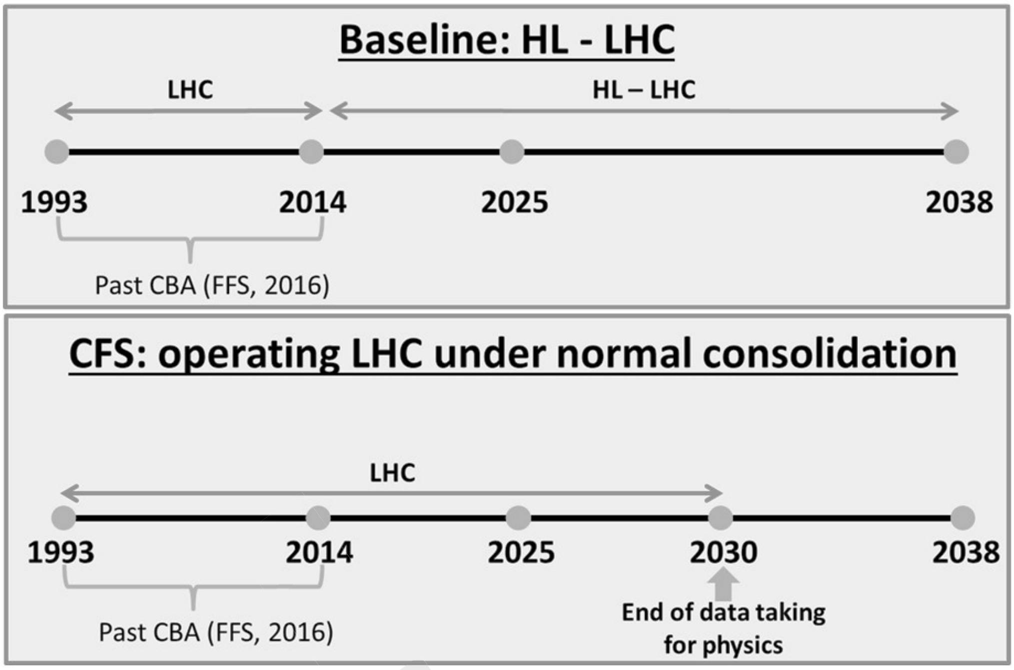
After having identified relevant costs and benefits in step (*i*), the analyst must decide how to apportion them over time. In the simplest case, when no prior information is available, a flat time profile can be used. In other cases, one might rely on sources of extraneous information (e.g. the time profile of costs and benefits of similar RI), or use selected insights from economic theory to design their time profile over the project horizon. In our analysis of the HL-LHC, we combine information collected at CERN and assumptions from previous studies.

The computation of costs and benefits for society often relies on complex and nonlinear formulas that involve a set of parameters, treated as random variables, and hence key ingredients of Monte Carlo simulations. Most of these parameters are unknown and must be estimated, and additional parameters are required to obtain a parametric representation of their distribution. The statistical distribution selected for a given parameter of the CBA model must yield reasonable results. For instance, enforcing non-negativity of costs and benefits requires selecting truncated distributions. Once a distribution has been selected, the values of its parameters must be set and different solutions can be adopted. The parameters can be estimated from the data, selected from previous studies or elicited from experts' opinions. For instance, in the case of operating and investment expenditures, alternative scenarios are often provided in the technical design report of the RI. These scenarios typically include a reference case - that can be used as a measure of the centre of the distribution – an optimistic and a pessimistic alternative that work as bounds for the support of the distribution.

4. Social costs and benefits of the HL-LHC project

For the deterministic CBA of the HL-LHC project we compute the *NPV* as the difference between the reference projection (i.e. HL-LHC) and a CF scenario entailing the operation of the LHC without the HL upgrade until the end of its life.

Figure 1. Timeline of the CBA: reference and counterfactual scenarios



Source: authors based on discussion with CERN experts.

This is determined by the depletion of the LHC discovery potential in the absence of a significant luminosity upgrade. In such a scenario, after 2031 experimental activity ends and CERN staff is relocated. Only planned maintenance and repair activities are considered from this point onwards. After the collider is switched off, the equipment left in the tunnel and the underground infrastructure would be subject to appropriate monitoring and safety procedures without being operated. A minimum of cooling, ventilation, electricity, and water supply would remain operational, as no decommissioning is planned.

Since HL-LHC is an upgrade of the LHC, the horizon of our analysis is $t = 1993, \ldots, 2038$ (see Brüning and Rossi, 2019 for the RI's life cycle) and encompasses the LHC program before the start of the HL project in 2015. Construction works for the LHC started in 1993, the year that marks the beginning time-period of our analysis. Therefore, costs and benefits from 1993 to 2014 coincide with those considered by Florio et al. (2016) in the CBA of the LHC, while 2015 marks the start of the HL-LHC program as far as our CBA is concerned. The first run of the HL-LHC is expected around

2026-2027. The operational phase of the HL-LHC is scheduled to last until 2038. Figure 1 provides a visual representation of the HL-LHC schedule.

Social costs and benefits of the HL-LHC are briefly discussed below, while details on definition, measurement and projections are in the Appendix.

Social Costs (*DC*). Social costs include operating and investment expenditures (*DF*) and indirect costs for society (*DX*). Negative externalities, *DX*, are assumed to be negligible since CERN is making several efforts to minimize externalities related to civil-engineering work.[19] It follows that the estimation of social costs boils down to measuring operating and investment expenditures, that is: $DC = DF$. Denoting with $t_0 = 2016$ the base year for discounting, such costs include past ($t < t_0$), current and future expenditures ($t \geq t_0$) that are attributed to the LHC accelerator complex programme and are paid by CERN and by the four major experiments (i.e. ATLAS, CMS, LHCb, ALICE).

V*alue of training* (*DH*). The value of training is proxied by the expected salary increase over the course of the career for those who have been involved as ESR in the project, relative to their peers elsewhere. This category includes technical students, doctoral students and post-doctoral researchers younger than 30 years who have been enrolled in a CERN education program and those who have been registered as "users" of CERN's experiments and are between 30 and 35 years old.

Earlier studies show that being part of CERN's activities as an ESR allows researchers to acquire cutting-edge knowledge and is a signal for future employers of their value, thus leading to better employment opportunities in the future, reflected in salary increases (Camporesi, 2001; Camporesi et al., 2017; Catalano et al., 2021).

---

[19] There are some examples pertaining to the HL-LHC. Part of the earth excavated to build new underground structures will be reused to create platforms for buildings and landscaping. The remaining portion of material will be taken to nearby waste storage facilities to limit the transport distance. A limited number of lorries transporting such wastes are constrained to travel during off-peak hours of working days. Constant environmental monitoring, the construction of temporary noise barriers and a project to landscape and plant the worksites with local tree species should also contribute to reduce externalities. See https://voisins.cern/en for further examples.

Technological spillovers (*DT*). This category includes two classes of benefits: (*i*) industrial spillovers, proxied by earning increases for companies working with CERN as suppliers and (*ii*) the estimated value of free software developed in the frame of the HL-LHC programme.

Increases in earnings derive from the commercialization of new products, the filing of patents, the discovery of new business opportunities, the development of more efficient operations and the gain of experience which arise as a consequence of the relationship with CERN (see, e.g., Åberg and Bengtson, 2015; Bianchi-Streit et al.,1984; Bastianin et al., 2021; Castelnovo et al., 2018; Florio et al., 2018b).

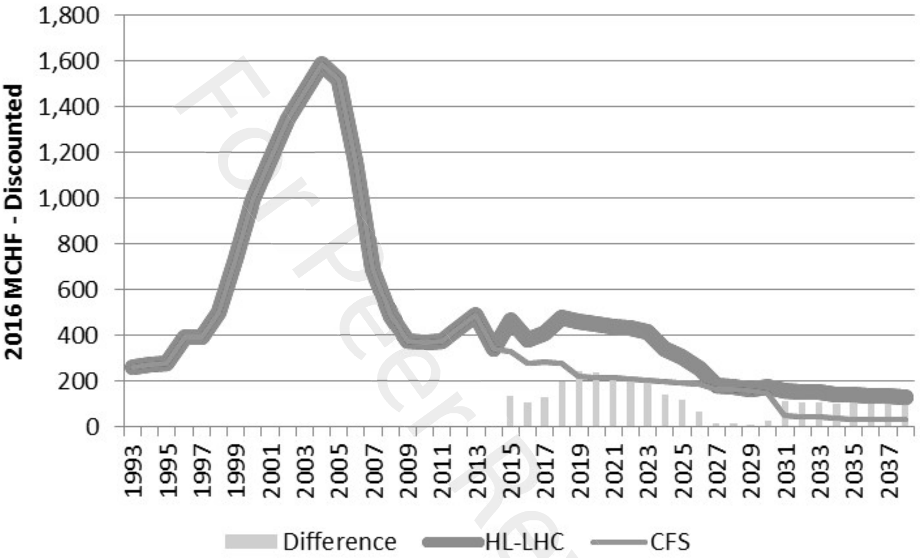
As for benefits related to the availability of free software, we multiply the cost of licences for alternative commercial software by an estimate of the number of users of two software packages made available for free and developed at CERN for the LHC program (ROOT and GEANT4).

C*ultural effects* (*DL*). For the sake of brevity, we list the categories of cultural effects that enter the estimate of *DL*, referring the interested reader to the Appendix for further details. The first – and most relevant – component quantifies the "tourism attractiveness" of CERN. Estimation of this category of cultural effects is implemented with the "travel cost method" (see Brown & Mendelsohn, 1984; Carson, 2012) and hinges on historical statistics about onsite CERN visitors and visitors of CERN travelling exhibitions. The remaining components quantify the engagement of users with media reporting about the HL-LHC program, HL-LHC-related websites and social media, volunteer computing programs and other media-related benefits such as movies and non-scientific books about CERN.

Publications and pre-prints (*DS*). Publications and pre-prints proxy the productivity and performance of scientists involved in HL-LHC activities and hence impact on their curricula and future work opportunities. The benefits of publications are measured in terms of the number of citations to papers produced at CERN for the HL-LHC program and consider their production cost relying on scientists' salary (Carrazza et al., 2016;).

Public good value (*DE*). The public good value represents a positive externality of new knowledge being generated for society at large. The magnitude of such externality is estimated by quantifying the willingness to pay for the creation of new knowledge by taxpayers. See Florio and Giffoni (2020) for an application to colliders at CERN and related literature in environmental and cultural economics.

Figure 2. Social costs (discounted 2016 MCHF)



Source: authors' calculations. The figure shows the discounted costs over the 1993-2038 period for HL-LHC (thick line) and CFS (thin line) and their difference (bars).

5. Results

5.1 CBA of the HL-LHC project

Figure 2 shows the evolution of discounted total costs over the 1993-2038 period. Up to 2015, the pattern of this variable reflects the baseline LHC program. In 2015, procurement for the HL upgrade starts and so do investment and operating costs for such programme. Comparison with the CF scenario reveals that, from 2015 onwards, the HL-LHC involves higher investment and operating costs, although in some periods, such as the end of 2020s, costs are expected to drop because the accelerator will not be operated during the so called “Long Shutdown 3” when the final components needed for the luminosity upgrade will be installed. Electricity expenditure explains a good share of the higher operating costs of the HL-LHC when compared with the CF scenario under which at the end of the 2020s scientific activity is expected to cease.

As reported in Sections 3 and 4, our CBA model considers several societal benefits. Focusing on the differential between the HL-LHC and CF scenarios,[20] Figure 3 shows the relative contribution of each category of benefits to the total. Effects related to the impact of CERN on human capital (*DH*) account for 40.30% of the total, while technological spillovers (*DT*) amount to about 37.70% of discounted total social benefits (*DB*). As for *DT*, “industrial spillovers” alone represent 29.20% of the increment in *DB*, while software account for an additional 8.50% of the total. Moving to the remaining categories of benefits, we see that the contribution to *DB* of cultural benefits (*DL*) and public good value (*DE*) is 5.65% and 10.73%, respectively.[21] The value of publications for scientist (*DS*), is 5.63%.

Given the importance of *DH* and *DT*, it is worth investigating these categories of social benefits more thoroughly. As for *DH*, panel (*a*) of Figure 4 illustrates why this social benefit is key for the positive evaluation of CERN’s programmes. In the absence of an upgrade to the LHC –the cornerstone of CERN’s scientific activity – ESR are expected to be either employed in alternative research programmes at CERN or working in other RIs outside CERN.[22] Figure 4(*b*) focuses on benefits for hi-tech suppliers, which represent the second most important contribution to total benefits. Without the luminosity upgrade, investment and operating costs will involve a progressively smaller share of hi-tech procurement and hence a smaller benefit for suppliers.

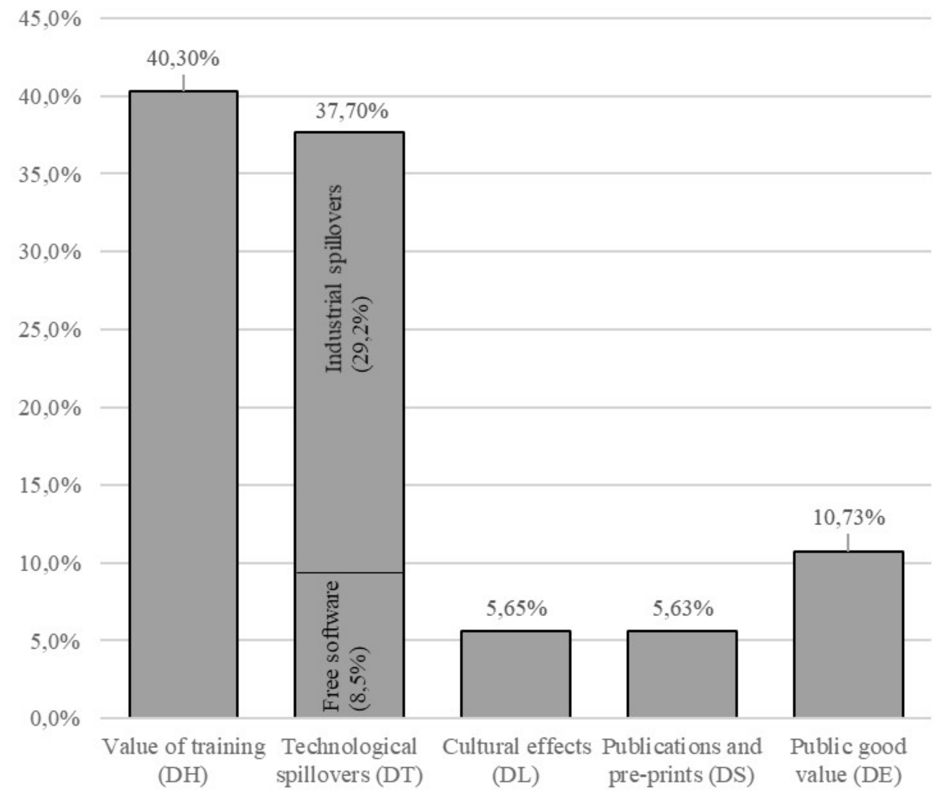
Figure 3. Percent contribution of different benefits to total social benefit

---

[20] For the remainder of this subsection, we drop subscript *j* since we are focusing on the differential between the *HL-LHC* and *CF* scenarios.

[21] As for *DL*, detailed analysis of the different items in this category reveals that benefits related with the “tourism attractiveness” of CERN are by far the largest and most important element.

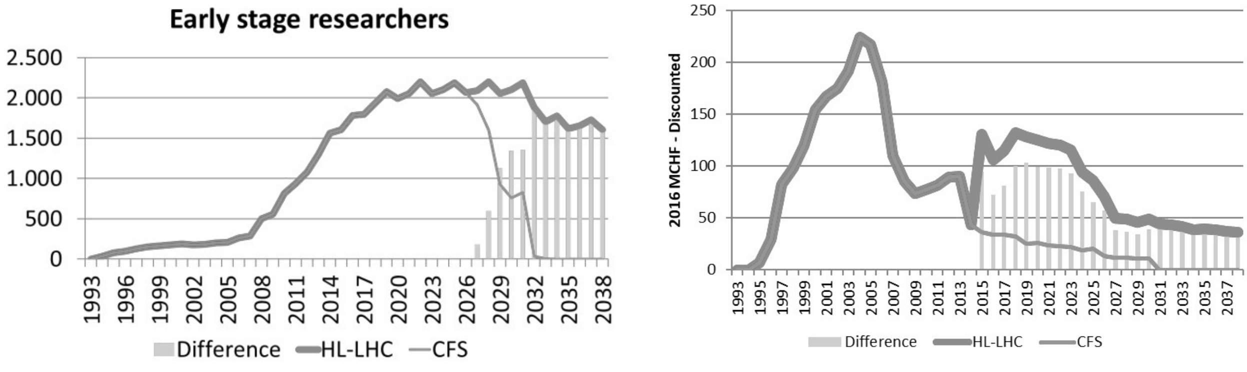
[22] One example is the Circular Electron Positron Collider (CEPC). This is circular collider – with a circumference of 80 km - first proposed by the Chinese high-energy physics community in 2012. See http://cepc.ihep.ac.cn/



Source: authors'calculations.

As a result of the CBA, when comparing the HL-LHC project with the CF scenario, the incremental NPV is 2219 MCHF, suggesting a positive contribution to society's well-being. Consistently, the benefit/cost ratio is 1.8. Both these indicators imply that investing in the HL upgrade yields a positive return for society.

Figure 4. Early Stage Researchers at CERN and benefits for hi-tech suppliers under different scenarios



(*a*) Number of ESR: 1993-2038 (*b*) Benefits for hi-tech suppliers: 1993-2038

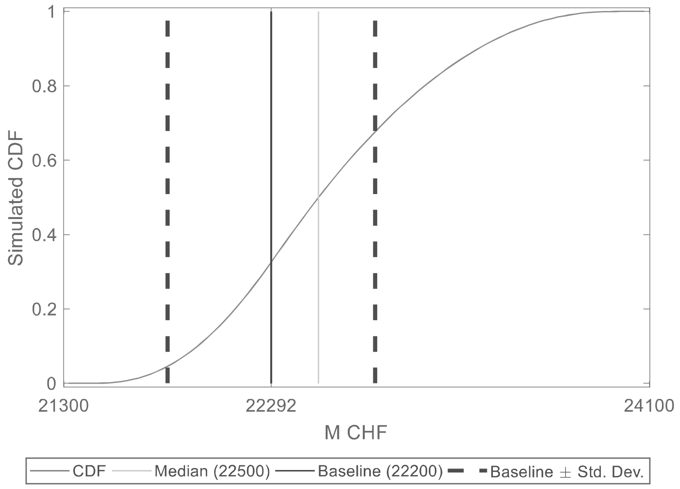
Notes: panel (a) shows the number of Early Stage Researchers over the 1993-2038 period for HL-LHC (thick line) and CFS (thin line). Panel (b) shows discounted benefits for hi-tech suppliers over the 1993-2038 period for HL-LHC (thick line), CFS (thin line) and their difference (bars).

5.2 Monte Carlo analysis of the HL-LHC project

The deterministic CBA provides evidence that the HL upgrade of the LHC creates value for society. These results, along with the fact that HL-LHC upgrade has already been financed by CERN Member States, led us to consider only this scenario in the probabilistic analysis. In other words, we illustrate the impact of risk in the case the decision to invest in the upgrade has been already taken.

We implement the Monte Carlo analysis of the HL-LHC scenario relying on 50,000 simulation rounds. At each run of simulations, we draw from the distribution of 15 different parameters underlying the CBA of the HL-LHC project. These parameters are the limits of the support, measures of location and - in some cases - of the scale of probability distributions of stochastic variables. For the sake of brevity, here we comment only on results for NPV, costs and selected benefits. The remaining results and further details are in the Appendix.

Figure 5. HL-LHC discounted costs: Monte Carlo results



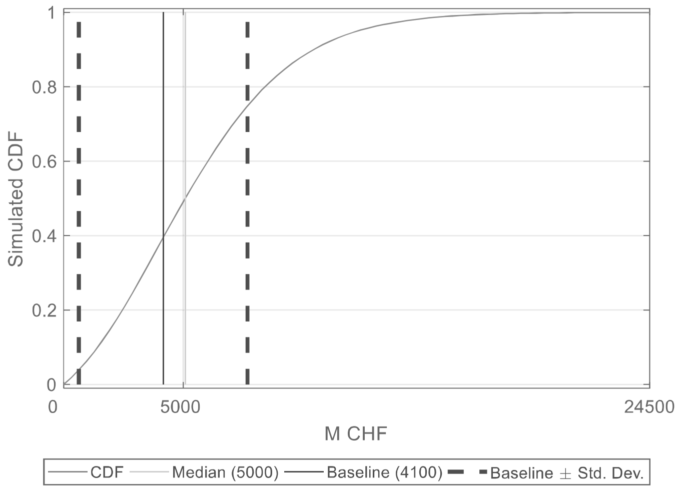
Average	22568
Median	22517
Min	21517
Max	23883
Std. Dev.	496

Notes: the figure shows the simulated Cumulative Distribution Function (CDF) based on 50,000 runs. We also show the baseline or reference value (i.e. the vertical black line; the value resulting from the CBA), the median simulated value (vertical grey line) and a 68% confidence interval (CI) for the baseline value (i.e. the area between the two vertical dashed lines). The table shows some descriptive statistics for the simulated values. Discrepancies due to numerical rounding

Monte Carlo simulation of the HL-LHC cost distribution relies on a triangular distribution for total costs and several categories of social benefits. The shape of the triangular distribution is defined by three parameters: the minimum, the maximum and the mode. To provide reasonable estimates of the bounds of the distribution of total costs, we rely on two scenarios for the evolution of HL-LHC costs provided by the CERN Finance Department. A pessimistic scenario implies a total cost that exceeds by 11% the reference value and an optimistic scenario that implies a 4% reduction of the reference total cost. The mode of the total cost distribution corresponds to the total cost of the HL-LHC used to produce the results in the previous section. Figure 5 shows that the median of the Monte Carlo distribution of $DC_{HL\text{-}LHC}$ is above the reference case, but is contained in the 68% confidence interval for the reference value.

We now turn to the analysis of the Monte Carlo distribution of benefits for hi-tech firms, one of the components of $DT_{HL\text{-}LHC}$. The undiscounted value of this benefit is computed, from the value of procurement ($PROC_t$) as follows: $PROC_t \times HT \times M \times SI$. In the Monte Carlo analysis, we consider the share of hi-tech procurement (HT), the sales multiplier (M) and the average sales increase (SI) as random variables and draw from their distribution. The sales multiplier, M, is assumed to be triangularly distributed with support 1.4-4.2 based on the analyses of Bianchi-Streit et al. (1984) and Schmied (1987). The modal and reference value of M is set to 3. SI is assumed to follow a Normal distribution with mean equal to 13% and standard deviation equal to 10% calibrated on data for a sample of CERN suppliers. This assumption implies that simulated values of SI can be negative, representing losses for the collaborating firm. To avoid applying a multiplier to losses, we truncate the distribution of SI to zero so that benefits for collaborating hi-tech firms cannot be negative. Figure 6 highlights that the distribution of discounted benefits for hi-tech suppliers of the HL-LHC program has a long right tail, however the 68% confidence interval around the reference value has a lower bound that lies close to zero, suggesting that the magnitude of this category of social benefits is highly uncertain. To the extent that benefits for hi-tech suppliers represent a key ingredient for passing the CBA test, CERN should continue investing in knowledge transfer activities – such as those documented by Nielsen and Anelli (2016) - so as to boost spillovers for its suppliers.

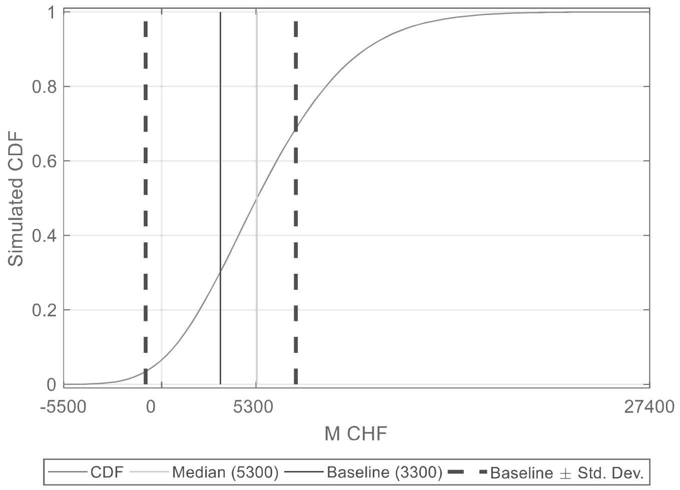
Figure 6. Benefits for hi-tech suppliers: Monte Carlo results



Industrial spillovers	
Average	5055
Median	4621
Min	0
Max	24404
Std. Dev.	3733

Notes: see notes to Figure 5.

Figure 7. HL-LHC: Distribution Function of the Net Present Value



NPV	
Average	1757
Median	1489
Min	-15844
Max	24314
Std. Dev.	4637

Notes: see notes to Figure 5.

Lastly, we focus on the simulated distribution of the NPV (Figure 7). The 68% confidence interval about the reference value does contain zero and hence we cannot rule out the possibility of observing a zero or even negative NPV. However, the probability of a negative NPV is about 6.3% with an estimated Monte Carlo error of 0.012.

5.2.1 Conditional scenario: a cost overrun

We now illustrate how the Monte Carlo analysis can be also used to construct conditional scenarios useful to verify the importance of specific assumptions on variables used in the CBA. In this case, we build a pessimistic conditional scenario focusing on costs, however one could also imagine alternative formulations that focus on other policy-relevant cases (e.g. the attractiveness of CERN for researchers or the impact of an important scientific discovery on the number of publications). As mentioned earlier, a cost overrun is one of the main sources of risk for Big Science projects and RIs. To this end, we keep the discounted value of costs at the maximum value, that is 11% higher than the reference value (i.e. at the upper bound of the triangular distribution used earlier). This exercise is as a very rough approximation of a scenario involving a rise in extra costs for whatever reason. We then run the Monte Carlo simulation conditioning on that value for costs. This exercise shows that even under the pessimistic scenario about costs, the probability of a positive NPV remains relatively high at 80%.

6. Concluding remarks

This paper illustrates how a probabilistic CBA model of large-scale RI can be implemented with a Monte Carlo approach. Moreover, we have applied such methodology to the CBA of the HL- LHC. This is the upgrade of the LHC, worth about one billion CHF, the social benefits of which are expected to be in excess of costs even ignoring any future scientific discovery. We now discuss how our research contributes to the literature on the ex-ante assessment of the socioeconomic impact of Big Science, usually involving projections over a time horizon of several decades.

The main question in the study of investment in science is that decisions must be taken in the present to discover unknown features of nature. This makes investment projects in this field particularly risky. Any large-scale investment, even those in traditional sectors such as transport or energy, are prone to errors when formulating scenarios, for example about future evolution of service demand or costs, but these errors are usually about the intensity, not about the existence of demand itself. Science is about generating knowledge based on testing theories, and, especially at the frontier, a theory may lead to nowhere. Only observation and experiments may prove or disprove a theory, and this often requires research activities conducted with costly and specialized equipment over long time periods.

Given such intrinsic uncertainty about the outcome of scientific investigation, a social CBA of Big Science projects was, until recently (Florio and Sirtori 2016), considered impossible. As a consequence, the forecast of the unknown was considered unfeasible, thus leaving investment decisions to politicians and scientists, supported only by qualitative studies of potential socio-economic impacts. Our approach suggests that constructing projection scenarios is indeed feasible, as long as the researcher focuses on all the critical variables of a Big Science project for which a probability distribution can be estimated, while setting aside the others. The result is only a partial view of the future, nevertheless a useful one as it leads to the understanding that what can be predicted may be sufficient to unveil the net social benefits of investing in science with today's perspective. Perhaps the project will not discover what is it was designed to discover (gravitational waves, for example, were measured 50 years after the first experiments to measure them began) but its implementation triggers a cascade of impacts that, to a certain extent, can be predicted. These impacts include those on scientific outputs, human capital creation, technological spillovers and cultural values. Assigning probabilities to such impacts is possible, and can be done in its simplest form by defining a range between a maximum and minimum value and the shape of a function, usually requiring just three parameters. Implementing a Monte Carlo simulation of the project's net present value conditional to such underlying distribution is feasible, and a Monte Carlo error can be minimised by increasing the number of iterations.

A crucial point is that the distributions need not be very precise if the researchers are able to collect the information on a relatively high number of critical variables (e.g 10-20), However, to collect such information expert advice is needed, typically by scientists and managers of RIs. This can be done by repeated sessions where the interviewee is required, first, to confirm that the variable is well defined (e.g. expected number of PhD students per year, number of orders to technological suppliers, number of scientific papers); second to provide the boundary values; third to select a distribution from a menu of alternatives. In certain cases, this crucial process may be managed as a focus-group or as a Delphi exercise (although this would require, preferably, third party experts). Thus our approach is complementary rather than a substitute for other strategies building scenarios.

The resulting NPV probability distribution is a dynamic concept, as it captures expectations on the future conditional to the information set of the present, and in principle can be updated as new information becomes available.

In our case study, nobody knows if the HL-LHC will discover new physics beyond the Standard Model, but we can quantitatively assess the probability of its socio-economic impact with a measurable statistical confidence. While any Big Science project is unique, our approach ideally could be replicated to several other projects in different scientific fields to further explore methodological improvements, particularly in the mechanism of elicitation of expert information for the design of the probabilistic model.

References

Abada, A. et al (2019). FCC-*ee*: the lepton collider. *The European Physical Journal Special Topics*, 228(2), 261-623.

Åberg, S., & Bengtson, A. (2015). Does CERN procurement result in innovation? *Innovation - The European Journal of Social Science Research*, 28(3), 360–383.

Autio, E., Hameri, A. P., & Vuola, O. (2004). A framework of industrial knowledge spillovers in big-science centers. *Research Policy*, 33(1), 107-126.

Bastianin, A., Castelnovo, P., Florio, M., & Giunta, A. (2021). Big science and innovation: gestation lag from procurement to patents for CERN suppliers. *Journal of Technology Transfer*, 1-25.

Bastianin, A., & Del Bo, C. F. (2020). Public procurement in Big Science: politics or technology? The case of CERN. *Industrial and Corporate Change*, forthcoming.

Bastianin, A., & Florio, M. (2019). Initial guidelines for a social cost-benefit analysis of the FCC programme (No. CERN-ACC-2019-0037). FCC-DRAFT-MGMT-2019-002.

Battistoni, G., Genco, M., Marsilio, M., Pancotti, C., Rossi, S., & Vignetti, S. (2016). Cost–benefit analysis of applied research infrastructure. Evidence from health care. *Technological Forecasting and Social Change*, 112, 79-91.

Bianchi-Streit, M., Blackburne, N., Budde, R., Reitz, H., Sagnell, B., Schmied, H. and Schorr, B. (1984), Economic utility resulting from CERN contracts (Second Study). CERN, 84-14, https://cds.cern.ch (last accessed March 2019).

Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2017). *Cost-benefit analysis: concepts and practice*. Cambridge University Press.

Brent, R. J. (2011). An implicit price of a DALY for use in a cost-benefit analysis of ARVs. Applied Economics, 43(11), 1413-1421.

Brown Jr, G., & Mendelsohn, R. (1984). The hedonic travel cost method. *Review of Economics and Statistics*, 66(3), 427-433.

Brüning, O., & Rossi, L. (2019). The high-luminosity large hadron collider. *Nature Reviews Physics*, 1(4), 241-243.

Brüning, O., & Rossi, L. (Eds.). (2015). *The High Luminosity Large Hadron Collider: the new machine for illuminating the mysteries of Universe* (Vol. 24). World Scientific Publishing Company.

Camporesi, T. (2001). High-energy physics as a career springboard. *European Journal of Physics*, 22(2), 139.

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.

Carrazza, S., Ferrara, A., & Salini, S. (2016). Research infrastructures in the LHC era: a scientometric approach. *Technological Forecasting and Social Change*, 112, 121-133.

Carson, R. T. (2012). Contingent valuation: A practical alternative when prices aren't available. *Journal of Economic Perspectives*, 26(4), 27-42.

Castelnovo, P., Florio, M., Forte, S., Rossi, L., & Sirtori, E. (2018). The economic impact of technological procurement for large-scale research infrastructures: Evidence from the Large Hadron Collider at CERN. *Research Policy*, 47(9), 1853-1867.

Catalano, G., Giffoni, F., & Morretta, V. (2021). Human and social capital accumulation within research infrastructures: The case of CERN. *Annals of Public and Cooperative Economics*. https://doi.org/10.1111/apce.12317

CERN (2020). *CERN Quick Facts*. Available online at (last accessed: May 2021): https://home.cern/resources/brochure/cern/quick-facts

Colby, E. R., & Len, L. K. (2016). Roadmap to the future. In, Chao, W. Chou (Eds.), *Reviews of Accelerator Science and Technology: Volume 9: Technology and Applications of Advanced Accelerator Concepts*. Chapter 1. World Scientific Publishing Co.

Congressional Budget Office, (2007). The Uncertainty of Budget Projections: A Discussion of Data and Methods. Available online at: https://www.cbo.gov/sites/default/files/cbofiles/ftpdocs/78xx/doc7837/03-05-uncertain.pdf

Del Bo, C. F. (2016). The rate of return to investment in R&D: The case of research infrastructures. *Technological Forecasting and Social Change*, 112, 26-37.

Del Bo, C. F., Florio, M., & Forte, S. (2016). The social impact of research infrastructures at the frontier of science and technology: The case of particle accelerators. *Technological Forecasting and Social Change*, 100(112), 1-3.

Dennis, M. A. (2017). Big Science. Encyclopedia Britannica. https://www.britannica.com/science/Big-Science-science

ESFRI - The European Strategy Forum on Research Infrastructures, 2016. *Strategy report on research infrastructures*. Available online at: http://www.esfri.eu/roadmap-2016 (last accessed: February 2021)

European Commission (2014). *Guide to Cost-Benefit Analysis of Investment projects*. DG Regional and Urban Policy.

European Commission (2017). *Horizon 2020 – Work Programme 2018–2020. European research infrastructures (including e–Infrastructures)*. Annex 4, European Commission Decision C(2017)7124 of 27 October 2017. Available online at: https://ec.europa.eu/program.

European Commission (2021). *Economic Appraisal Vademecum 2021-2027:General Principles and Sector Applications*. Directorate-General for Regional and Urban Policy. Available online at: https://ec.europa.eu/regional_policy/sources/docgener/guides/vademecum_2127/vademecum_2127_en.pdf

Evans, L. (2016). Particle accelerators at CERN: From the early days to the LHC and beyond. *Technological Forecasting and Social Change*, 112, 4-12.

Florio, M. (2014). *Applied welfare economics: Cost-benefit analysis of projects and policies*. Routledge.

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

Florio, M., Bastianin, A., & Castelnovo, P. (2018a). The socio-economic impact of a breakthrough in the particle accelerators' technology: a research agenda. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 909, 21-26.

Florio, M., Forte, S., & 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.

Florio, M., & Giffoni, F. (2020). A contingent valuation experiment about future particle accelerators at CERN. *PloS one*, 15(3), e0229885.

Florio, M., Giffoni, F., Giunta, A., & Sirtori, E. (2018b). Big science, learning, and innovation: evidence from CERN procurement. *Industrial and Corporate Change*, 27(5), 915-936.

Florio, M., & Sirtori, E. (2016). Social benefits and costs of large scale research infrastructures. *Technological Forecasting and Social Change*, 112, 65-78.

Gaillard, M. K., Grannis, P. D., & Sciulli, F. J. (1999). The standard model of particle physics. *Reviews of Modern Physics*, 71(2), S96.

Giffoni, F., Schubert, T., Kroll, H., Zenker, A., Griniece, E., Gulyas, O. Angelis, J., Reid, A., Vignetti, S. (2021) *State of play – Literature review*. RIPATHS deliverables no. T3.2.

Granger, C. W., & Jeon, Y. (2007). Long-term forecasting and evaluation. *International Journal of Forecasting*, 23(4), 539-551.

Greenberg, D., Walter, J., & Knight, G. (2013). A cost-benefit analysis of the random assignment UK Employment Retention and Advancement Demonstration. *Applied Economics*, 45(31), 4335-4354.

Griliches, Z. (1958). Research costs and social returns: Hybrid corn and related innovations. *Journal of Political Economy*, 66(5), 419-431.

Griliches, Z. (1992). The search for R&D spillovers. *Scandinavian Journal of Economics*, S29-S47.

Helmers, C., & Overman, H. G. (2017). My precious! The location and diffusion of scientific research: evidence from the Synchrotron Diamond Light Source. *The Economic Journal*, 127(604), 2006-2040.

Johansson, P. O., & Kriström, B. (2018). *Cost-benefit analysis*. In Cambridge Elements of Public Economics, Cambridge University Press.

Jones, C. I., & Williams, J. C. (1998). Measuring the social return to R&D. *Quarterly Journal of Economics*, 113(4), 1119-1135.

Knight, F. H. (1921) *Risk, Uncertainty, and Profit*. Boston, MA: Hart, Schaffner & Marx; Houghton Mifflin Company Lawrence, M., Goodwin, P., O'Connor, M., & Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. International Journal of forecasting, 22(3), 493-518.

Long J.F., Mcmillen D.B. (1987) A survey of Census Bureau population projection methods. *Climatic Change*. 11(1-2):141-77.

Lee, R. (2011). The outlook for population growth. *Science*, 333(6042), 569-573.

Linton, J. D. (2008). Why big science has trouble finding big money and small science has difficulties finding small money. *Technovation*, 28(12), 799-801.

Maresova, P., Sobeslav, V., & Krejcar, O. (2017). Cost–benefit analysis–evaluation model of cloud computing deployment for use in companies. *Applied Economics*, 49(6), 521-533.

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

Müller, U. K., & Watson, M. W. (2016). Measuring uncertainty about long-run predictions. *Review of Economic Studies*, 83(4), 1711-1740.

Nelson, R. R., & Romer, P. M. (1996). Science, economic growth, and public policy. *Challenge*, 39(1), 9-21.

Nilsen, V., & Anelli, G. (2016). Knowledge transfer at CERN. *Technological Forecasting and Social Change*, 112, 113-120.

OECD, 2014. *The impacts of large research infrastructures on economic innovation and on society: case studies at CERN*. Available online at: https://www.oecd.org/sti/inno/CERN-case-studies.pdf.

Önkal, D., Sayım, K. Z., & Gönül, M. S. (2013). Scenarios as channels of forecast advice. *Technological Forecasting and Social Change*, 80(4), 772-788.

Pflaumer, P. (1988). Confidence intervals for population projections based on Monte Carlo methods. *International Journal of Forecasting*, 4(1), 135-142.

Raftery, A. E., & Ševčíková, H. (2021). Probabilistic population forecasting: Short to very long-term. *International Journal of Forecasting*, forthcoming.

Riordan, M., Hoddeson, L., & Kolb, A. W. (2015). *Tunnel visions: the rise and fall of the Superconducting Super Collider*. University of Chicago Press.

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

Schmied, H. (1987), About the quantification of the economic impact of public investments into scientific research, *International Journal of Technology Management*, 2(5–6), 711–729.

Scaringella, L., & Chanaron, J. J. (2016). Grenoble-GIANT Territorial Innovation Models: Are investments in research infrastructures worthwhile? *Technological Forecasting and Social Change*, 112, 92-101.

Vuola, O., & Hameri, A. P. (2006). Mutually benefiting joint innovation process between industry and big-science. *Technovation*, 26(1), 3-12.

Appendix to "Projecting the socio-economic impact of a Big Science Center: the world's largest particle accelerator at CERN"

Andrea Bastianin[a,*] Chiara F. Del Bo[a] Massimo Florio[a] Francesco Giffoni[b]

This draft: October 12, 2022

[a] Department of Economics, Management and Quantitative Methods, University of Milan.

[b] CSIL, Centre for Industrial Studies.

[*] *Corresponding author*: Department of Economics, Management and Quantitative Methods, University of Milan. Via Conservatorio, 7. I-20122 - Milan – Italy.
Email: andrea.bastianin@unimi.it

A1. A CBA Model for Big Science projects

The expected Net Present Value (*NPV*) of Big Science project "*j*" over a given time horizon:

$$E(NPV_j) \equiv E[DB_j - DC_j] \tag{A1}$$

DB_j and DC_j are the cumulative sum of discounted social costs and benefits of the project over its lifespan. Social costs include not only operating and investment expenditures (DF_j), but also indirect costs and externalities for the society (DX_j).

Following Florio and Sirtori (2016), in the case of Big Science projects we can identify the following set of social benefits:

- the value of training, or human capital formation, for students and early-stage researchers (DH_j);
- technological spillovers for collaborating firms and other economic agents (DT_j);
- cultural effects for the public (DL_j);
- the value of academic publications and pre-prints for scientists (DS_j);
- the existence or public good value of the RI (DE_j).

We stress that this set of benefits might not be exhaustive and that, if not appropriate, the analyst needs not consider the whole set. Assuming that all of the previous social benefits are relevant, we can rewrite Equation (A1) as:

$$E(NPV_j) \equiv E[DB_j - DC_j] = E[(DH_j + DT_j + DL_j + DS_j + DE_j) - (DF_j + DX_j)] \tag{A2}$$

A2. CBA of the HL-LHC project: implementation

This section details the implementation of the CBA for the HL-LHC project, the counterfactual (CF) scenario and the assumptions of the Monte Carlo exercise summarized in Table A1.

A2.1 Discounting social costs and benefits

Discounted social costs and benefits are computed setting a base period, t_0, and relying on a social discount rate, r. The social discount rate, r, used in the CBA of investment projects, reflects the opportunity cost of capital from an inter-temporal perspective for society as a whole (i.e. it reflects the social view of how future benefits and costs are to be valued against present ones) and does not necessarily coincide with any of the interest rates set in financial markets. See Drupp et al. (2018) for a recent analysis of the topic. Discounted social costs at time t are given by: $DC_{j,t} = C_{j,t} /(1 + r)^{\tau}$ where $\tau \equiv t - t_0$ for $t = 0, \ldots, H$ and H is the time-horizon relevant for the CBA. Cumulative discounted costs are thus given by: $DC_j = \sum_t DC_{j,t}$. Cumulative discounted benefits are obtained in the same fashion. Notice that, since the base period t_0 does not necessarily coincide with the starting date of the project ($t = 0$), social costs and benefits are discounted whenever $t > t_0$ ($\tau > 0$), while they are compounded as $t < t_0$ ($\tau < 0$).

In our application of the CBA method to the HL-LHC project, the base year for social discounting and inflation adjustments is $t_0 = 2016$. The analysis is carried out in million Swiss Francs (MCHF) at constant prices and the social discount rate is set at $r = 3\%$ as recommended by the European Commission (2014). After the base year (i.e. $t > t_0$) we assume constant prices.

A2.2 Social Costs

Both the HL-LHC and CF scenarios represent extensions of the LHC program, therefore for the 1993-2014 period, costs and benefits are the same as those documented by Florio et al. (2016). Over the 2015-38 period, we obtained cost estimates for both scenarios from CERN's Finance and Administrative Processes Department. An assumption that will be maintained in our analysis involves scientific personnel cost. We exclude the cost of scientific personnel at CERN from total operating costs. The rationale for not considering this cost item is that we assume this share of cost to be balanced by the "production cost" of scientific publications, that is one of the benefits that is imputed to RI. Since we account for this benefit as a separate category, we remove it from costs to avoid issues of double counting. To subtract the share of personnel cost from overall expenditures, we sourced data on different cost items from the "Final Budget of the Organization for the sixty-second financial year 2016". For each cost item, we have computed the percentage of costs to be attributed to the LHC programme (i.e. this includes the HL-LHC upgrade). Scientific personnel is assumed to 32% of the total personnel cost of CERN and is assumed to be 100% cost for the experimental collaborations.

A2.3 Social benefits

A2.3.1 Value of training (*DH*)

DH is the social benefit that early stage researchers (ESR) - who spent a period working at the HL-LHC project - are expected to enjoy in terms of career development when entering the labour market after their years of service at CERN. Many studies provide empirical evidence of the existence of such a benefit stemming from the acquisition of scientific and technical skills as well as communication competences during the experience at CERN (Bianchin at al., 2019). The value of training is estimated focusing on ESR's salary evolution over time as compared to their peers without the CERN experience (Camporesi et al., 2017; Catalano et al., 2021).

To quantify the value of training, we applied the following formula (Florio et al., 2016):

$$H_t = N_t \times YL \times W \times WP \qquad \text{(A3)}$$

where:

- N_t is number of ESR that each year leave CERN to enter the labour market. Over the period 1993-2038, 40,567 ESR are expected to work at the HL-LHC project (20,770 in the CF scenario). We assume that N_t has a triangular distribution centred at the baseline yearly value for each category of ESR (i.e. Fellows, Users, and PhD Students and Post-Doctoral students) and with limits set at ±15% from this baseline value.
- *YL* is the expected duration of working life that we set equal to 42 years.

- *W*, the average salary after CERN, is used to compute the salary premium (see *WP* below). Average salary depends on the years of working experience, the level of education, and the sectors of employment of ESR (industry, academia, research centres, and other sectors, including finance, public administration and no-profit). See Florio et al. (2016) for details.
- *WP* is the "salary premium", that is the expected increment in salary due to experience at CERN that we assume to be triangularly distributed with mode 11.8%, minimum 11.3% and maximum 12.3% (see Florio et al., 2016; Camporesi et al., 2017; Catalano et al., 2021).

Notice that this benefit extends to *YL* years after the last cohort of ESR leaves CERN. It follows that *DH* denotes the discounted value of training over this timespan.

A2.3.2 Technological spillovers (*DT*)

Technological spillovers arising from CERN scientific programmes can be divided into two broad classes: (*i*) industrial spillovers, proxied by earning increases for companies working with CERN as suppliers (*ii*) the estimated value of free software developed within the frame of the HL-LHC programme.

Industrial spillovers. Collaborative procurement relations between CERN and its industrial suppliers improve suppliers' performance in terms of increasing sales in other markets and therefore their profit. The existence of a "CERN sales multiplier" associated with the CERN hi-tech procurement has been documented in numerous studies (see e.g., Schmied, 1977, 1987; Bianchi-Streit et al., 1984; Autio at al., 2004; Castelnovo et al., 2018). Accordingly, we assumed that this benefit is proportional to the value of hi-tech HL-LHC procurement contracts and employed the following formula for its calculation:

$$PROC_t \times HT \times M \times SI \qquad \text{(A4)}$$

Where:

- $PROC_t$ is the total yearly value of procurement contracts for the HL-LHC programme. The total value of CERN procurement by year is sourced from the CERN Procurement Department.
- *HT* is the share of the procurement (over total procurement) classified as "hi-tech". We classified HL-LHC contracts as hi- and lo-tech as follows. Each order was assigned a value on a five point scale by experts at CERN. This scale is as follows: (1) off-the-shelf orders with low technological intensity; (2) off-the-shelf orders with an average technological intensity; (3) mostly off-the-shelf orders but requiring some careful specifications; (4) hi-tech orders with a moderate to high specification activity intensity to customize product for the HL-LHC programme; (5) products at the frontier of technology with an intensive customization work and co-design involving CERN staff. Orders with technological intensity equal or greater than 3 are assigned the "hi-tech" label (Florio at al., 2016). The share of hi-tech procurement represents 85% of the experiment collaborations and 35% for the other CERN procurement orders. In the Monte Carlo, we assume that the share of hi-tech procurement for the collaborations to have a triangular distribution with mode equal to 58% and limits set at 55% and 90%. As for CERN procurement, the parameters of the triangular distribution are 35% (mode), 34% (minimum) and 75% (maximum). According to CERN staff procurement of experimental collaborations has typically higher technological intensity than CERN's orders.
- *M* is the sales multiplier. Based on existing studies (see e.g. Schmied, 1977, 1987; Bianchi-Streit et al., 1984), its baseline value was set at 3 and assigned a triangular distribution with support in the range 1.4 – 4.2.
- *SI* is the average incremental profit for HL-LHC suppliers. Profits are estimated relying on the EBITDA (i.e. Earnings Before Interest, Tax, Depreciation and Amortization) of a sample of CERN suppliers. *SI* is assumed to be Normally distributed with the mean equal to 13% and standard deviation equals to 10% (Castelnovo et al., 2018). The distribution is truncated at zero to avoid negative values.

Estimated value of free software developed in the frame of the HL-LHC programme. This benefit is driven by software and ICT developments needed for the HL-LHC programme and made available for free outside the high energy physics (HEP) community. Following Florio et al. (2016) and Catalano at al. (2019) we estimate this class of benefits focusing on two open-source software packages: ROOT (a data analysis platform written in C++ used also in the finance sector; see https://root.cern/) and GEANT4 (i.e. a toolkit for the simulation of the passage of particles through matter used not only in physics, but also in medical and space sciences; see https://geant4.web.cern.ch/node/1).

We thus estimated the opportunity cost for the purchase of an equivalent commercial software in the case of ROOT and the cost required for development of an analogous tool for what concerns GEANT4. We thus proceed as follows

$$[R\_USERS_t \times (P\_ALT \times Y\_USE)] + [G\_ORG_t \times G\_AC] \qquad \text{(A5)}$$

Where:

- R_USERS_t is the yearly number of users of the ROOT. In 2013 non-CERN ROOT users outside the HEP community were estimated to be about 25,000 worldwide and forecasted to be 55,000 in 2025 and assumed to remain at that level up to 2038 (Florio at al., 2016). The number of new ROOT users by year up to 2038 is thus obtained assuming a constant growth rate. We rely on a triangular distribution with a range of ±20% with respect to the baseline case (i.e. the number of users in 2038). The opportunity cost of ROOT represents the cost for buying an alternative commercial software (e.g. Matlab, SPSS, Stata,…). It is computed as $P\_ALT \times Y\_USE$, where P_ALT is the market price of a comparable commercial software and Y_USE is the number of years before such commercial licences become obsolete. As a baseline value of a commercial single-user licence, P_ALT, we set CHF 1,754 per year. A triangular distribution is used for P_ALT with range in CHF 1,170-2,339 and mode equal to the baseline value. As for Y_USE, based on interviews at CERN, we opted for a trapezoidal probability distribution function. Its parameters are the two vertexes equal to 3 (first vertex) and 10 (second vertex), the minimum equal to 0 and the maximum set at 20. The baseline value was set at 7.5 years.

- G_ORG_t is the yearly number of research centres, space agencies and firms in which GEANT4 is routinely used. As of 2016, 38 centres contributed in some form to the development of the code and 12 centres were passive users (i.e. did not actively contributed at all). The opportunity cost of GEANT4 is estimated to be about CHF 35 million. CERN contributed for about 50% of such development cost. Accordingly, the avoided cost for the aforementioned 38 centres was reduced by a half and amount to about CHF 460,000 per centre. For the remaining 12 centres, the full development cost. In the Monte Carlo exercise, we rely on a baseline value for G_AC that takes into account the cost of GEANT4 for all these centres. We then rely on a triangular distribution with a range of ±30% from the baseline value (see Florio et al., 2016 for additional details).

As highlighted in a number of interviews with the staff of CERN, the HL-LHC programme will need additional ICT developments. We then also included a third category of benefits labelled "Importance of new ICT" to capture further positive externalities deriving from the development of additional software and services for storage (e.g. scalable file system services) and computing solutions (e.g. Cloud and Grid computing management tools) for the HL upgrade (see Di Meglio et al., 2017). However, given the difficulties to quantify the impact of such ICT technologies, we rely on a triangular distribution that scales the discounted value of benefits produced by ROOT (see Bastianin and Florio, 2018). Such distribution is centred at 2 and has support in the range 1-45.

A2.3.3 Cultural effects (*DL*)

The cultural impact of the HL-LHC on the general public is estimated focusing on visitors of CERN, its travelling exhibitions, websites, and outreach activities. Florio et al., (2016) showed that for the

LHC the largest share of this benefit (about 60%) is generated by the touristic attractiveness of the LHC machine, and specifically by onsite CERN visitors and visitors of CERN travelling exhibitions around the world. In contrast, the share of the cultural benefit associated with the reach of media reporting on LHC - including websites the users of LHC-related social media (e.g. YouTube; Twitter; Facebook; Google+) and other media-related benefits such as movies and non-scientific book - is marginal at best. Accordingly, in the case of HL-LHC the focus was on the benefit to onsite CERN visitors and visitors of CERN travelling exhibitions. We and it was calculated by applying the Travel Cost Method as follows (see Champ et al., 2015):

$$(OV_t \times TRC) + (TV_t \times TE\_WTP) \tag{A6}$$

Where:

- OV_t is the number of onsite CERN visitors. Based on CERN data, Florio et al., (2016) estimated that from 2004 to 2025 the total number of CERN visitors was about 1.7 million people (78,000 per year). Visitors include both people visiting the experiment facilities and the permanent CERN exhibitions (i.e. Microcosm and Universe of Particles in the Globe of Science and Innovation). From 2026-2038 we keep the yearly number of visitors constant at the level of the 2004-2025 period.
- *TRC* stands for the travel cost. It was estimated to be about CHF 700 per visitor in the baseline scenario (2016 prices). It includes the cost of travel to CERN, accommodation and meals during the visit. The *TRC* is depends on several variables, including the country of origin of the visitors (the cost increases as the distance from CERN increases), the mode of transport (plane, train, etc.), the length of stay, and the value of travellers' time based on HEATCO data.
- TV_t is the number of visitors to CERN travelling exhibitions. According to CERN it was between 30,000 and 70,000 for the period 2006-2013. We assumed a constant number of 40,000 visitors per year from 2014 to 2038.

- *TE_WTP* stands willingness-to-pay (WTP) for the "travelling exhibitions" and reflects the price visitors would pay to access the CERN exhibition around the world. This WTP was assumed to be CHF 1 per visitor (assuming local transport).

- For the LHC (1993-2025), Florio et al., (2016) estimated a total discounted cultural benefit (*DL*) equal to CHF 1.1 million (2016 prices). It was CHF 57,000 in 2025. In the scope of this study, we assumed that for the 2026-2038 period there is no increase of visits at CERN due to the HL-LHC. Hence, we kept the undiscounted value of *DL* constant up to 2038. We treated, however, this benefit as a stochastic variable by assigning it a triangular distribution with a range of ±20% with respect to the baseline value.

The share of cultural effects related to visitors of CERN HL-LHC experiment Websites and users of (HL-) LHC-related social media, they are worth respectively 12% and 2% of the total amount of cultural benefits (see Bastianin and Florio, 2018; Florio at al., 2016 for details).

A2.3.4 Publications and pre-prints (*DS*)

Given that we track cites to HL-LHC publications and pre-prints this benefit extends beyond the lifetime of the programme. It is worth pointing out that these benefits relate to the quantity of publications and citations, not to their impact on the scientific community. More precisely, in the computation of *DS* we assume that the publications by scientists working at the HL-LHC program have a production cost – captured by the cost of scientific personnel costs – that is offset by their value. This implies that neither the cost, nor the benefit of publications authored by HL-LHC scientists – denoted as L_0 - enter the CBA. We thus estimate *DS* considering only articles and pre-prints (L_1) by scientists not working at the HL-LHC program that cite L_0 papers. We thus estimate *DS* relying on the production cost of L_1 publications and the value of citations and downloads of L_0 and L_1 publications (Ferrara and Salini, 2012).[1]

In the period 1993-2025, 22,886 L_0 papers were produced by LHC scientists, with a peak of 2,008 publications in 2013 - five years after the discovery of the Higgs Boson in 2008 (Florio at al., 2016). We applied the marginal cost approach to the estimation of this benefit. The marginal social value of a single publication is thus proxied by its marginal production cost of producing that is proportional to the salary of scientists (European Commission, 2014). Florio at al., (2016) estimated that the unit (marginal) production cost of a paper authored by LHC scientists – denoted as L_0 – was about CHF 12,876 on average in 2016 prices.[2] Given that, by definition, the value of a L_0 paper is proportional to its production cost, in the CBA framework, the benefits of L_0 publications cancel out with their cost of production represented by the scientific personnel cost. Therefore, we excluded the value of L_0 publications in the computation of this benefit and only considered the value of the first-wave citing papers - denoted as L_1 - i.e. of papers by scientists who are not involved in the (HL-) LHC program and cite L_0 papers, and the value of these citations. We also considered the value of their citations in subsequent papers, i.e. the second-wave citing papers – denoted as L_2.

To compute the value of HL-LHC scientific production, we proceed as follows:

$$VL_{1,\,t} + Cit\_VL_{1,\,t} + Cit\_VL_{2,\,t} \tag{A7}$$

Where:

- $VL_{1,t}$ is the yearly value of L_1 papers. First, the number of L_1 papers, which is a function of the number of L_0 publications. Based on the bibliometric model described by Carrazza et al. (2016) the number of L_1 papers to be considered was estimated to be 241,671 until 2050 (i.e. 11 L_1 papers for each L_0 publication on average). The second element entering the estimation of $VL_{1,t}$ is the unit marginal cost of L_1 papers. It is estimated to be CHF 367.95 (in 2016 prices)

---

[1] It is important to note that the benefit does not measure the value of the content of the publication, i.e. *the knowledge per se*, which is unpredictable from an ex-ante perspective (e.g. the future value of potential applications theoretically described and predicted by the publication), but its impact on the scientific community (see Florio and Sirtori, 2016).

[2] It corresponds to EUR 11,011 in 2013 prices. The marginal production cost was calculated as a function of the LHC scientists' annual salary, the share of scientists' time dedicated to research, and the number of co-authors.

per L_1 paper and can be interpreted as the value of the contribution of L_0 publications to L_1 papers.[3].

- $Cit\_VL_{1,t}$ is the yearly value of L_1 citations to L_0 papers. It depends on two variables: the number of L_1 papers as above and the value of the citation. According to the marginal cost approach, this value represents the opportunity cost of the time employed by a researcher to read the cited paper. The citation unit marginal cost was set to CHF 116 per L_1 paper.
- $Cit\_VL_{2,t}$ is the yearly value of the citations of L_1 papers by L_2 papers. It depends on the number of L_2 papers, which is a function of the number of L_1 publications. Our bibliometric model predicted 862,100 L_2 papers until 2050 or 4 citations per L_1 paper. The value of the single citation then computed as above.

To sum up, the benefit from scientific production consists of three components: (*i*) the value of L_1 papers; (*ii*) the value of L_1 citations to L_0 papers; and (*iii*) the value of L_2 citations to L_1 papers.

For the HL-LHC, the value of academic publications and pre-prints (*DS*) is extended until 2063 to consider the life-cycle of scientific publications and citations beyond the lifetime of the particle collider programme. It was estimated to be CHF 613 million (discounted in 2016 prices) in the whole period 1993 – 2063 and it was worth 2% of the total social benefits (see Bastianin and Florio, 2018 for further details). For the Monte Carlo analysis, we assume that in 2031 there might be a spike in the number of HL-LHC papers that is comparable to what happened in 2013 after the discovery of the Higgs Boson. The existence of such spike and its magnitude is highly uncertain and depends on the scientific discoveries arising from the operation of the HL-LHC. In the Monte Carlo analysis we rely on a triangular distribution that scales the number of citations in HL-LHC papers in 2031. The distribution has mode at 1 (meaning that the number of publications in 2031 is equal to that in 2013) and has range in 0.1 – 1.

---

[3] The value depends on the number of references in L_1 publications. On average, it was 35 per paper.

A2.3.5 Public Good Value (*DE*)

CERN produces knowledge advances about the nature and the origins of the universe, which can be considered a “public good”. The production knowledge is financed by the contributions of its Member States (MS) and therefore taxpayers are ultimately the funders of CERN’s investment projects. While there is no immediate application or use in view of this knowledge for the public, citizens however, may be willing to financially support CERN research for the pleasure and the utility that something could be discovered. Therefore, the public good value represents the benefits due to the fact that new knowledge might be generated for the society, and in particular for non-users of the HL-LHC (i.e. people who currently do not directly use the services of the HL-LHC, but are better-off simply because they know that new knowledge might be created) See Florio and Giffoni (2020).

The public good value was estimated based on the taxpayers’ “willingness to pay” (WTP) for a particle collider research infrastructure:

$$NT_t \times WTP \tag{A8}$$

Where:

- NT_t is the number of non-users interested by the calculation of this benefit. Based on evidence from field experiments, about 87.6 million non-users from both CERN MS and CERN non-MS in the period 1993-2038 were considered[4]. See Florio and Giffoni (2020) and Florio et al. (2020). The *WTP* is assumed to have a triangular distribution with mode at CHF 1.5 and range in CHF 0.1-2.

---

[4] This number was calculated by considering 73% of the number of 18-74 years old people with at least a tertiary education coming from CERN Member States. Experiments based on surveys show that that 27% of respondents has a zero WTP and that having an university degree is an important determinant of the WTP for research. In addition to CERN MS and based on evidence on the flow of visitors at CERN, it was assumed that people from non-MS had to be included in the calculation as well. Accordingly, 21% of 18-74 years old people with at least a tertiary education coming from CERN non-MS was accounted for (see Florio at al., 2016 for further details).

A3. Additional Tables

#	Variable		Category	Distribution	Baseline	Parameters[(a)]				Unit	Source
1	Discounted total cost	$DC_{HL\text{-}LHC}$	Social Costs	Triangular	Discounted total cost	0%	-4%	11%	-	% difference from baseline	CERN Finance Department
2	ESR leaving CERN for the labor market	N_t	Value of training (*DH*)	Triangular	Annual value for each category (i.e. Fellows, Users, and PhD Students and Post-Doctoral students)	0%	-15%	15%	-	% difference from baseline	CERN staff
3	Salary premium	*WP*	Value of training (*DH*)	Triangular	11.8%	11.8%	11.3%	12.3%	-	% increase in annual salary	Florio et al. (2016); Camporesi et al. (2017); Catalano et al. (2021).
4	Share of hi-tech procurement for CERN	*HT*	Technological Spillovers (*DT*)	Triangular	35%	35%	34%	75%	-	% of hi-tech procurement over total procurement for the HL-LHC	Florio et al. (2016). Discussion with CERN staff
5	Share of hi-tech procurement for experimental collaborations	*HT*	Technological Spillovers (*DT*)	Triangular	58%	58%	55%	90%	-	% of hi-tech procurement over total procurement for the HL-LHC	Florio et al. (2016). Discussion with CERN staff
6	Sales Multiplier	*M*	Technological Spillovers (*DT*)	Triangular	3	3	1.4	4.2	-	Sales Multiplier	Schmied (1977, 1987); Bianchi-Streit et al. (1984)
7	Average profit increase	*SI*	Technological Spillovers (*DT*)	Normal truncated at zero (i.e. only non-negative values)	13%	13%	10%	-	-	%	Castelnovo et al. (2018)
8	ROOT users	R_USERS_t	Technological Spillovers (*DT*)	Triangular	Root users in 2038	0%	-20%	20%	-	% difference from baseline	Own assumption based on Florio et al. (2016)
9	Price of alternative software	*P_ALT*	Technological Spillovers (*DT*)	Triangular	1754	1754	1170	2339	-	CHF per person per year	Own assumption based on the cost of commercial software licenses
10	Years before alternative software becomes obsolete	*Y_USE*	Technological Spillovers (*DT*)	Trapezoidal distribution	7.5	0	3	10	20	years	Discussion with CERN staff.
11	Avoided development costs for GEANT4	*G_AC*	Technological Spillovers (*DT*)	Triangular	Cost of GEANT4	0%	-30%	30%	-	% difference from baseline	Florio et al. (2016)
12	Importance of new ICT	-	Technological Spillovers (*DT*)	Triangular	Baseline benefits associated with ROOT	2	1	45	-	CHF (multiplicative factors)	Di Meglio et al. (2017); Bastianin and Florio (2018); CERN staff.
13	Total Cultural Benefits	*DL*	Cultural effects (*DL*)	Triangular	Baseline value	0%	-20%	20%	-	% difference from baseline	Own assumption
14	Number of publications in 2031	*DS*	Publications and pre-prints (*DS*)	Triangular	Publications in 2013	1	0,1	1	-	Scale factor for the baseline	Own assumption
15	Willingness to pay	*WTP*	Public good value (*DE*).	Triangular	1.5	1.5	0.1	2	-	CHF	Florio et al. (2020); Florio and Giffoni (2020)

(a) The triangular distribution has three parameters mode, minimum and maximum (in this order in the table); the Trapezoidal distribution has 5 parameters minimum, first vertex, second vertex and maximum.

Table A2. Cost Benefit Analysis of HL-LHC, Counterfactual (CF) scenarios, cost and benefit differential

Discounted MCHF 2016	HL-LHC	%	CF	%	Difference	%
Total cost (DH_j)	22292		19356		2936	
Total Benefit (DB_j)	25608		20453		5155	
Human Capital (DH_j)	8379	33%	6302	31%	2077	40%
Technological Spillovers (DT_j)	10187	40%	8244	40%	1943	38%
- Software	*6029*	*24%*	*5591*	*27%*	*438*	*9%*
- Hi-tech Suppliers	*4158*	*16%*	*2653*	*13%*	*1505*	*29%*
Cultural Effects (DL_j)	3319	13%	3028	15%	291	6%
Publications (DS_j)	613	2%	322	2%	290	6%
Public Good Value (DE_j)	3110	12%	2557	12%	553	11%
NPV	3316		1097		2219	
Benefit/Cost ratio (DB_j / DC_j)	1.15		1.06		1.76	

Notes: columns 3, 5, 7 report the percent contribution of each category to total benefit, DB_j. Column 6 shows, for each item, the difference between HL-LHC and the counterfactual (CF) scenario. NPV is the net present value, that is the difference between discounted social benefits (DB_j) and costs (DC_j). Benefits are given by: $DB_j = DH_j + DT_j + DS_j + DL_j + DE_j$. Technological spillovers, DT_j, have two components: "Software" and "Hi-tech Suppliers". Benefit/cost ratio is given by: DB_j / DC_j where j = HL-LHC, CF, Difference.

References

Autio, E., A. P. Hameri and O. Vuola (2004), ‘A framework of industrial knowledge spillovers in big-science centers,’ *Research Policy*, 33(1), 107–126.

Bastianin, A., & Florio, M. (2019). Initial guidelines for a social cost-benefit analysis of the FCC programme (No. CERN-ACC-2019-0037). FCC-DRAFT-MGMT-2019-002.

Bianchi-Streit, M., Blackburne, N., Budde, R., Reitz, H., Sagnell, B., Schmied, H. and Schorr, B. (1984), Economic utility resulting from CERN contracts (Second Study). CERN, 84-14, https://cds.cern.ch (last accessed March 2019).

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.

Camporesi, T. (2001). High-energy physics as a career springboard. *European Journal of Physics*, 22(2), 139.

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.

Carrazza, S., Ferrara, A., & Salini, S. (2016). Research infrastructures in the LHC era: a scientometric approach. *Technological Forecasting and Social Change*, 112, 121-133.

Castelnovo, P., Florio, M., Forte, S., Rossi, L., & Sirtori, E. (2018). The economic impact of technological procurement for large-scale research infrastructures: Evidence from the Large Hadron Collider at CERN. *Research Policy*, 47(9), 1853-1867.

Catalano G., Carrazza S., Castelnovo P. and Florio M. (2019), Assessing the value of CERN’s free and open source software: the case of ROOT

Catalano, G., Giffoni, F., & Morretta, V. (2021). Human and social capital accumulation within research infrastructures: The case of CERN. *Annals of Public and Cooperative Economics*. https://doi.org/10.1111/apce.12317

Champ, P. A., Boyle, K. J., & Brown, T. C. (Eds.). (2015). A Primer on Nonmarket Valuation (2nd ed.). Springer.

Di Meglio, A., Girone, M., Purcell, A. & Rademakers, F. (2017) “CERN openlab white paper on future ICT challenges in scientific research”, https://cds.cern.ch/record/2301895

Drupp, M. A., Freeman, M. C., Groom, B., & Nesje, F. (2018). Discounting disentangled. *American Economic Journal: Economic Policy*, 10(4), 109-34.

European Commission (2014). *Guide to Cost-Benefit Analysis of Investment projects*. DG Regional and Urban Policy.

Ferrara, A., & Salini, S. (2012). Ten challenges in modeling bibliographic data for bibliometric analysis. *Scientometrics*, 93(3), 765-785.

Florio, M., Forte, S., & 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.

Florio, M., & Giffoni, F. (2020). A contingent valuation experiment about future particle accelerators at CERN. *PloS one*, 15(3), e0229885.

Florio, M., Giffoni, F., and Catalano, G. (2020). 'Should governments fund basic science? Evidence from a willingness-to-pay experiment in five universities', *Journal of Economic Policy Reform*, vol. 23(1), pp. 16-33 https://doi.org/10.1080/17487870.2018.1547638

Florio, M., & Sirtori, E. (2016). Social benefits and costs of large scale research infrastructures. *Technological Forecasting and Social Change*, 112, 65-78.

Schmied, H. (1977). A study of economic utility resulting from CERN contracts. *IEEE Transactions on Engineering Management*, 4, 125–138.

Schmied, H. (1987), 'About the quantification of the economic impact of public investments into scientific research,' *International Journal of Technology Management*, 2(5–6), 711–729.