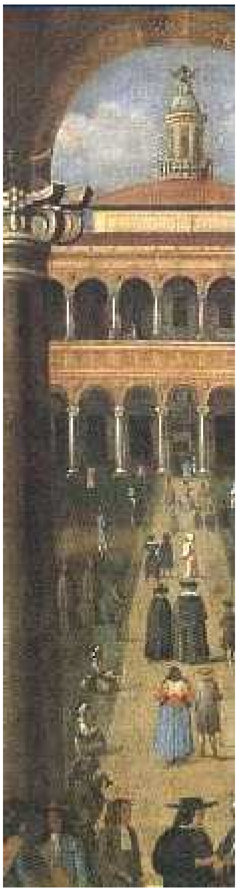


Milan
European
Economy
Workshops

THE EVALUATION OF RESEARCH INFRASTRUCTURES: A COST-BENEFIT ANALYSIS FRAMEWORK

MASSIMO FLORIO **EMANUELA SIRTORI**

Working Paper n. 2014-10
DICEMBRE 2014

JEAN MONNET CHAIR
Economics of European Integration

UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI ECONOMIA, MANAGEMENT E METODI QUANTITATIVI

Via Conservatorio 7
20122 Milano
tel. ++39 02 503 21501 (21522) - fax ++39 02 503 21450 (21505)
http://www.economia.unimi.it
E Mail: dipeco@unimi.it

XI Milan European Economic Workshop
Università degli Studi di Milano, 22-23 giugno 2012

The Evaluation of Research Infrastructures: a Cost-Benefit Analysis Framework

Massimo Florio[1] and Emanuela Sirtori[2]

[1] *Department of Economics, Management and Quantitative Methods, University of Milan*

[2] *CSIL Centre for Industrial Studies*

This draft: 21.10.2014

Abstract

When decision-makers consider pure and applied research infrastructures, such as genomics platforms, astronomic observatories, nanoelectronic laboratories, oceanographic vessels, or particle accelerator facilities (just to mention some examples) are faced by this question: what is the net social benefit of these costly scientific ventures and of the public goods they produce? The answer is often given qualitatively, or even rhetorically, by scientists and other stakeholders in these projects. But can we go beyond anecdotal evidence, narratives and ad hoc studies and try a structured ex-ante and ex-post evaluation of the socio-economic impact of research infrastructures? This paper explores some of the methodological issues involved in a CBA framework for capital-intensive scientific projects. The paper proposes a conceptual model based on the estimation of quantities and shadow prices of cost aggregates, and of six main categories of economic benefits (pure value of discovery, knowledge outputs, technological spillovers, human capital formation, cultural effects and services to third parties). Empirical approaches are suggested for further applied research, including the use of probability distribution functions to generate expected net present values of research infrastructures by Monte Carlo methods.

Keywords: Research infrastructures, Cost-benefit analysis, Externalities, Public good, Knowledge

JEL Codes: D61, D81, I23, O32

Acknowledgments: The paper has been produced in the frame of the research project '*Cost/Benefit Analysis in the Research, Development and Innovation Sector*' sponsored by the EIB University Research Sponsorship programme (EIBURS), whose financial support is gratefully acknowledged. Further details on this research project can be found at: http://www.eiburs.unimi.it/

The authors are grateful to participants at the Milan European Economy Workshop 2014 June 12-13 "The Cost-Benefit Analysis of Research Infrastructures" and particularly to Per-Olov Johansson and Mark Mawhinney for their helpful comments on a previous version, and to Donatella Cheri and Chiara Pancotti for research assistance.

European
Investment
Bank · Institute

1 Introduction

The number of capital-intensive research projects, e.g. in the field of genomics, astronomy, space exploration, high energy physics, chemistry of new materials, bio-molecular engineering, and several other fields, witnessed a substantial growth in the past decades. This growth was driven by two factors. On the one side, science evolves fast and more and more sophisticated and powerful experimental instruments need to be designed and constructed in order to back such a progress and push forward the frontiers of knowledge. On the other side, research is increasingly put at the centre of political agendas as a tool to stimulate economic growth.[1] In the European Union (EU), the 'Europe 2020'[2] overarching strategy for promoting dynamic and sustainable growth throughout the Member States includes the Innovation Union flagship initiative, aimed at transforming Europe into a world-class science performer, by establishing a common European Research Area and completing or launching the construction of priority European research infrastructures. The United States (US) and Japan, and more recently China, and other countries, are planning large-scale scientific ventures for the next decades.

The size of capital-intensive research projects has also never been greater. The term 'Big Science' was coined many years ago to describe the large-scale character and complexity of modern science, in contrast with the formerly predominant 'Little Science' (de Solla Price, 1963 and Weinberg, 1967). Major examples of Big Science are the International Space Station, the Hubble Space Telescope and the last generations of particle accelerators and colliders. Large scale scientific projects tend to entail substantial investment costs related to the design and construction of research infrastructures, which often rise, even considerably, from the ex-ante estimates, thus attracting numerous criticisms. The Superconducting Super Collider can be mentioned in this regard. Planned to be built in Texas with an initial estimated budget of USD 4.4 billion,[3] this 87 km circumference particle accelerator started to be built in the early Nineties. After having already spent USD 2 billion and dug 23.5 km of underground tunnel and 17 pits, the cost for the project completion rapidly surged to USD 11 billion and the project was eventually abandoned by the US Congress (Giudice, 2010; Baggott, 2012; Maiani, 2012).

The increasing number of research infrastructures financed and their increasing average cost call for a serious examination about whether it is worth or not to spend considerable amounts of money by governments, and ultimately by citizens in their capacity as tax-payers. Some researchers are highly optimistic about the value of scientific research promoted by Governments (e.g. Mazzucato, 2011[4]), but others are more sceptical (e.g. see Broad's review, 1990[5]). In principle, when a decision about investment priorities needs to be taken, costs of any project have to be assessed against the associated social benefits in order to check whether the latter exceed the former. Hence, the fundamental question is: how to guess the social net benefit of research infrastructures? The answer is often given qualitatively, or even rhetorically by scientists and the other stakeholders with an interest in the project implementation. Typically, the decision of funding highly expensive facilities, usually without any

---

[1] The role of R&D investments as a driving force for capital accumulation and, ultimately, long-run growth has been studied in the economic literature for example by Griliches (1980), Adams (1990) and Romer (1990). See a companion paper on the returns of capital investment in R&D infrastructures (Del Bo, 2014).

[2] European Commission Communication 'Europe 2020. A strategy for smart, sustainable and inclusive growth', COM(2010) 2020 final.

[3] Nominal prices.

[4] '[…] the role of the government, in the most successful economies, has gone way beyond creating the right infrastructure and setting the rules. It is a leading agent in achieving the type of innovative breakthroughs that allow companies, and economies, to grow, not just by creating the 'conditions' that enable innovation. Rather the state can proactively create strategy around a new high growth area before the potential is understood by the business community (from the internet to nanotechnology), funding the most uncertain phase of the research that the private sector is too risk-averse to engage with, seeking and commissioning further developments, and often even overseeing the commercialisation process. In this sense it has played an important entrepreneurial role.' (Mazzucato, 2011: 18-19).

[5] In his article, Broad (1990) presents the major criticisms against large and costly research ventures, spanning from their negative impact on the governmental budget, particularly in situations of nation's fiscal troubles, to the displacement of funds away from smaller and more promising research projects.

chance of earning a financial return, is supported by a coalition of scientists with the task of convincing funders of the project usefulness and legitimacy.[6] But can we go beyond that lobbying approach, and try a robust forecast or at least an ex-post evaluation of costs and benefits of research infrastructures?

Cost-Benefit Analysis (CBA) is a framework largely adopted by international institutions (e.g. the European Commission, the European Investment Bank and the World Bank) and governments in public decision-making to assess the socio-economic profitability of investment projects in other fields. CBA consists in assessing whether benefits accrued from project implementation are in excess of the estimated socio-economic costs, thereby showing if the project represents a net benefit to the whole society. The key strength of CBA is that it produces information of the project's net contribution to the society, synthesised into simple indicators, such as the Net Present Value (NPV).

Whatever the difficulty in estimating the social cost of any investment, because of lack of data or specific conceptual issues, particularly when externalities are considered,[7] a standard CBA theory for the estimation of their value to society is available and well established (see e.g. Drèze and Stern 1987, Florio 2014a, Johansson 1991, Pearce *et al.* 2006). In particular, there is a long worldwide experience in the social cost-benefit analysis of infrastructures in transport, energy or water, and more recently in environmental services, health, education, culture and other fields.

Some preliminary attempts to provide specific indications for consultants and public officers involved in appraising research infrastructure projects have been made. In 2009 the Czech government, in collaboration with the JASPERS team of the European Investment Bank, developed a working document providing guidance on the methodology to compute the CBA economic indicators for projects in this area. A revised and extended staff working papers has then been drafted by JASPERS (2013). Some empirical attempts to measure the economic return of investments in the research and development sector have been made at micro level, generally returning a positive result (see Del Bo, 2014 and Pancotti *et al.* 2014 for an overview of the literature). The new edition of the European Commission Guide to Cost-Benefit Analysis of Investment Projects (European Commission, forthcoming) has a new chapter on Research, Development and Innovation infrastructures. It is recognised, however, that the application of a CBA framework of analysis to the assessment of this typology of infrastructures still requires a sound theoretical framework to be developed. Thus our research question is straightforward: how should a CBA model for research infrastructures be designed?

This paper explores some of the methodological issues involved when evaluating capital-intensive research projects (both ex-ante and ex-post) through the CBA framework, and concludes that such a framework can be designed and tested empirically, with due caution given its experimental nature.

The framework can be used to answer the question of how benefits produced by large-scale research infrastructures ('Big Science') compare to their costs. In principle, the same model can be flexibly applied to assess costs and benefits of both pure and applied research infrastructures, regardless its size, even if some effects (particularly wider cultural effects and human capital accumulation) become relevant only when a certain critical mass is reached. The proposed CBA model applies to RIs operating in all scientific fields. Some examples recalled in this paper pertain to research infrastructures for physics, because we are currently testing the model by case studies of particle accelerators (see www.eiburs.unimi.it); however, the scope of the paper is more general. Moreover, the paper adopts both an ex-post and ex-ante project evaluation perspective.

---

[6] Or occasionally failing to convince them, as the above mentioned story of the Superconducting Super Collider shows: approved by President Reagan in 1987, it was then cancelled by the US Congress under Clinton in 1993 (Maiani, 2012).

[7] Projects aimed at tackling climate change are an extreme example. See the Stern Review (HM Treasury, 2006).

The structure of the paper is the following: in *Section two*, after defining the object of analysis, i.e. what we mean by research infrastructure, we outline the conceptual CBA model for research infrastructures and we propose and justify a taxonomy of benefits for this typology of infrastructural investments. *Section three* examines the social demand for the infrastructure and the social value of six main types of benefits. We discuss knowledge outputs, technological externalities, human capital development, wider cultural effects and services to third parties; we then suggest a possible way to think about the pure intrinsic value of discovery. For each of these six effects we mention empirical approaches for estimation of marginal social values. *Section four* introduces the treatment of the considerable uncertainty surrounding the CBA estimates: we suggest using risk analysis based on a Bayesian approach, involving probability distribution functions and Monte Carlo techniques. *Section five* concludes by putting together the cost and the benefit sides of the discussion, with some initial implications for decision making.

2 Definitions, hypotheses and conceptual framework

There is no established and agreed definition of research infrastructures (RIs). Different definitions and classifications are proposed in the literature and in policy documents, often pointing to large investment costs. The Strategy Report on Research Infrastructures prepared by the European Strategy Forum on Research Infrastructure (ESFRI, 2011) also includes among research infrastructures electronic surveys, such as the European Social Survey,[8] and other facilities for data collection and storage, like databases, archives, libraries and computer grids. Given the wide variety of facilities and instruments that are generally referred to in this field, as a starting point it is deemed necessary to provide a definition of infrastructures to which the CBA conceptual model developed in this paper can be applicable.

A research infrastructure is, after all, an infrastructure because it provides services to users. Thus it is helpful to review some definitions of this concept in economics. An influential review paper by Gramlich (1994) points to a definition in terms of large capital-intensive investment often associated with natural monopolies.[9] Typical examples of economic infrastructures comprise assets and equipment providing transport services, telecommunications, water and sewer management, and energy provision. Other infrastructures which more recently captured the attention of economists are related to education, healthcare, culture. In this vein, we argue that Gramlich’s definition can be used in our context, and that research services are not *per se* of a completely different nature as compared e.g. to transport, environmental, education, cultural or health services.

The fact that the benefits of a transport, energy, communication, environmental project can be measured and valued by CBA techniques, while for the services of a research infrastructure there is still no accepted CBA method, should not blur the issue at stake. After all, some decades ago it was often maintained that investment in some sectors characterised by intangible outputs, for example having an impact on human health, could not be evaluated by CBA techniques (see e.g. Baum and Tolbert 1985), while this is an accepted practice today (see e.g. Viscusi and Aldy, 2003; the World Health Organization guidelines, 2006). Valuing an intangible good such as culture has also been considered impossible for long time, but recently some efforts have been made to explore the economic value of culture, using techniques which are compatible with the CBA framework, such as contingent valuation (see the report by the UK Department for Culture, Media and Sport, DCMS, 2010). Indeed there are some ingredients of research infrastructure that are peculiar to them, but several others are shared with other categories of infrastructures.

---

[8] The European Social Survey is a network established to develop, store and study long time series of data used to monitor and interpret changes in European social attitudes and behaviour patterns.

[9] See below for a definition of natural monopoly in this context.

A critical ingredient of any infrastructure is high capital intensity. Capital fixed expenditure overcomes operating costs and is a large fraction of the total present value of project cost. This is particularly true in Big Science, which is performed using some of the most expensive machines ever built. Perhaps the most expensive RI so far is the International Space Station, whose total costs are reported by the European Space Agency to be around USD 100 billion over a 30 year period, including development, assembly and running costs for ten years.[10] The approximate investment cost of other Big Science projects is shown in Table 1. On this ground, we are going to exclude from the definition of research infrastructures all surveys, since the service they provide – i.e. the collection and elaboration of data – is more labour, rather than capital, intensive.

Table 1 Cost of Big Science: some examples (Million EUR inflated to 2013 prices)

Research Infrastructure	**Description**	**Approximate total investment cost (M EUR)**	**Source**
Superconducting Super Collider – SSC (USA)	Particle accelerator with a planned ring circumference of 87 km. The project was cancelled in 1993.	13,460[11]	Giudice (2010)
Large Hadron Collider – LHC (CERN, Switzerland)	The world's largest and most powerful particle accelerator. It started operations in 2008.	7,230[12]	Maiani (2012) and CERN (2009)
National Ignition Facility (USA)	Laser-based inertial confinement fusion research facility, built between 1997 and 2008 and operational since 2009.	3,350	GAO (2000) and press release
Large Electron-Positron collider – LEP (CERN, Switzerland)	Electron-positron accelerator. Commissioned in 1989 and closed down in 2000, it was the predecessor of LHC.	1,730[13]	Schopper (2009)
Central European Institute of Technology – CEITEC (Czech Republic)	Centre of excellence conducting research in the field of life sciences, advanced materials and technologies. It is currently under construction.	310	Data provided by the EIB
Extreme Light Infrastructure – ELI (Hungary)	The world's highest power laser, currently under construction.	310	Data provided by the EIB

Source: Authors based on different cited sources

Fixed investment costs of Little Science infrastructure[14] also tend often to be larger that operating costs. However the gap reduces going from most sophisticated and innovative infrastructures to more standard and easily replicable facilities and equipment.

Data about relatively smaller RIs, gathered from the European Portal on Research Infrastructure (Riportal)[15] confirm that the smaller the investment cost, the smaller the average number of permanent scientific/engineering staff operating the RI, internal users (scientists directly working at the RI's experiments), external users (other scientists making use of the results obtained by the internal staff working on the machines) and trainees and students.[16]

---

[10] http://www.esa.int/Our_Activities/Human_Spaceflight/International_Space_Station/How_much_does_it_cost

[11] 1993 last cost estimate.

[12] Including in-kind provision.

[13] Excluding external personnel and in-kind provision.

[14] Examples include the Italian Laboratory for the Study of the Effects of the Radiation on Material for Space, the Finnish Centre of excellence in Environmental Health risk analysis or the Hungarian Cyclotron of Atomki that provides accelerated particles that can be used for nuclear physics studies and for radioactive isotope production for application purposes.

[15] http://www.riportal.eu/public/index.cfm?fuseaction=ri.search.

[16] Del Bo (2014) performs a more in-depth analysis of these data.

A second ingredient of the definition of any infrastructure is the long time horizon involved in both the cost side and the benefit side. Some research infrastructures are still operating today after decades since their construction. For example CERN accelerators built in the late Fifties (Proton Synchrotron) and in the Seventies (Super Proton Synchrotron) are still used as injectors of proton beams in the most recently built collider (LHC). The time span of benefits is also extremely long, as it is discussed below.

Third, ‘standard’ economic infrastructures are often associated with externalities and spillover effects: economic benefits of infrastructures are usually not appropriated by the owner of the infrastructure, and we shall show that this is a core feature of research infrastructures as well.

Fourth, the “natural monopoly” feature is typical of many research infrastructures, particularly the largest ones, as production cost of the research outputs are subadditive. Decreasing average costs associated to prevailing fixed capital costs lead to limited or no competition among the research infrastructures. In the research sector there are facilities of which there will be no more than one in the country or in the world because a second one would be too costly or because the number of users is not big enough (an example of unique infrastructure is the already mentioned International Space Station). The preferred arrangement in these cases usually entails wide collaborations of scientists using the same infrastructure and exploiting its features to perform different experiments, either over time or at the same time. The fact that there is no explicit market for the services of the RI does not change the tendency to concentration, but only its mechanisms (Irvin and Martin, 1984a). However, sometimes in Big Science the same research question could be answered in principle by more than one research infrastructure.[17] As a matter of fact, a certain degree of competition among different scientific teams is considered important in the field of research, not only for speeding up the pace to achieve a given objective or make a discovery, but also to validate the respective findings (Baggott, 2012).

This adds to the interest of evaluating the relative costs and benefits of competing projects. In this context, we argue that ‘research’ relates to all those activities which elaborate data and information for creating new knowledge. According to this criterion, RIs include both *pure* research infrastructures, carried out for the main purpose of increasing the understanding of fundamental scientific principles and producing new ideas, but also *applied* research facilities aimed at acquiring new knowledge directed to a practical purpose (e.g. creating a new compound, technology or product).[18] University research laboratories generally fall into this category.[19] RIs typically pursue their objective using material facilities and instruments, but some of them can provide their service electronically. Actually, infrastructures based on high performance Information and Communication Technologies (sometimes called e-infrastructures, or digital infrastructure) can also be considered research infrastructures as long as they are essential to make particularly complex computations and simulations out of any human being’s reach, thus actually producing new knowledge. Examples are supercomputers and grid computing, consisting of computer resources specifically developed for processing very large volumes of data and producing outputs for scientific use.

---

[17] A comparative assessment of the advantages and disadvantages of the CERN Large Electron-Positron (LEP) collider was conducted by Martin and Irvine (1984). This exercise shows that even very large and cutting-edge accelerators might have a number of rivals playing some competition (even in an oligopoly framework) among themselves.

[18] According to the same criterion, material and virtual archives, databases, networks and cloud computing that serve to collect, store and share already existing data, without processing them in any way so as to create new knowledge, are not considered here research infrastructures. This kind of equipment and facilities certainly contribute at improving research efficiency, by making available a larger amount of data in a shorter time, but they do not directly produce new knowledge. We do not consider networks for research and education as research infrastructures either. The same holds for catalogues of digitalised books, journals, newspapers and libraries.

[19] However, some university departments are not to be considered research infrastructures, but rather education facilities, as long as their main objective is knowledge dissemination.

Most of RIs are single-sited,[20] but there are also examples of geographically distributed facilities, such as grid computing systems or atmospheric measurement stations located in different areas and recording data which are then centrally studied.[21] Distributed RIs are such that they provide a unique service by means of data recorded and/or computed in facilities located in different areas. An organization or network of mutually-independent RIs, each one providing its service without depending on the service provided by another RI of the same network and with different unrelated research questions, is not accounted as a single distributed RI, but as a number of single-sited infrastructures. In such cases there may be, however, network externalities to be considered in the project's impact assessment.[22]

To sum up, for the purpose of the CBA conceptual framework suggested in this paper we understand research infrastructures as high-capital intensity and long-lasting facilities and equipment, typically operating in 'monopoly' or 'oligopoly' conditions, whose objective is to produce social benefits through the generation of new knowledge, either pure or applied.

The literature on the benefits stemming from research is wide. In some earlier literature,[23] many 'positive outcomes' of research projects are listed,[24] including:

- knowledge creation and dissemination, possibly measured by a conventional economic value of scientific papers, books published and other outputs, such as research contracts granted from public and private funders;
- technological development, measured by estimated the economic value of patents granted, licence deals, spinoffs created and technologies transferred;
- human capital creation, usually valued through the salary of Masters and Ph.D. students carrying out research at the RI;
- employment effects, measured by the average salary for newly created jobs both at the RI and in spin-offs;
- social capital creation, understood as the development and strengthening of 'productive' interactions, shared values and mutual trust within the scientific community;
- others, related for example to reputation and marketing effects on RI's suppliers, development of tourism activities nearby the RI, increase interest towards science among young people, and a broadly defined added value for society thanks to the responses to 'grand challenges' and 'big quests' of life produced by scientific research and discoveries.

However, most of these effects, and the way they are often valued in earlier literature are inconsistent or only partially consistent with the theoretical framework of social CBA.[25] For example, for the purpose of CBA employment of scientists and technicians is an input for the project, hence when

---

[20] Single sited RIs include particle collider facilities, telescopes, space stations, research vessels and aircrafts, science parks, space research centres, laser light facilities, microscopy facilities, institutes for radio astronomy, biobanks for medical and biological study, research reactors, geographical research observatories, botanic research institutes, museum's laboratories for zoology, botany, earth and human science and supercomputers.

[21] Other examples are seismographic stations and aquaculture and laboratory testing facilities.

[22] Note that this definition of distributed research infrastructure differs from the one recently provided by the OECD, according to which a distributed infrastructure is a network or multi-national association of geographically-separated organisational entities that jointly operate a set of independent research facilities. An example is the European Very Large Baseline Interferometry Network that is a collaboration of the major radio astronomical institutes of Europe, Asia and Africa (OECD, 2014).

[23] See Salter and Martin (2000), Hallonsten et al. (2004), SQW Consulting (2008), Czech Ministry of Education, Youth and Sport and JASPERS (2009), Science and Technology Facilities Council (2010), COST Office (2010), JASPERS (2013), Bach (2013).

[24] See a companion paper by Pancotti, Pellegrin and Vignetti (2014) for a more extensive discussion of possible beneficial effects of research infrastructures often mentioned in the literature and considered in the appraisal and selection process of RI projects.

[25] In Annex 1 the most established model for CBA in a general equilibrium setting, developed by Drèze and Stern (1987 and 1990), is briefly recalled.

valued it is a project cost, and not a benefit.[26] A patent associated to the building or operation of a RI is an externality only if, properly valued, generates a windfall gain for, e.g., the project's supplier. By definition, there is no externality if in fact the research leading to the patent has been 'paid' by the firm/institute through its own expenditures. Moreover, any new product developed by firms of the RI's supply chain thanks to the experience and skills gained on-the-job is often suggested to be valued through the firm's turnover, while in fact it is the incremental profit (incremental sales net of incremental costs of production) that should be considered. Training of young scientists is not a benefit either: the benefit is the future welfare change associated to the training against a counterfactual where such advanced training is not provided. Thus valuing the human capital creation through the actual or incremental wage of former RI-students is misleading if a proper counterfactual is not taken into account. As to the social capital accumulation arisen from interactions among scientists, particularly when they come from different countries, in a CBA perspective it is not a benefit *per se*, but as long as it serves for increasing productivity of scientists: hence, it should be reflected in the benefit of knowledge output.

We propose to consider a simpler and shorter (but more precise) taxonomy of benefits than what has been often proposed in the earlier literature. Out taxonomy is parsimonious, as in fact we do not want to multiply items, and conservative, as we do not want to exaggerate the social impact of research projects. Yet, we believe it captures the main direct and indirect impacts of research infrastructures, and is robust across different types of projects.

The fundamental CBA model for research infrastructures, on which the rest of the paper is built and further elaborates, is presented by equation (1):

$$ENPV_{RI} = ENPV_u + ENPV_n = (EPV_{B_u} - EPV_{C_u}) + EPV_{B_n}. \qquad (1)$$

The expected net present value – ENPV (the expectation operator 'E' will be discussed later, and it will be dropped below to simplify notation[27]) of research infrastructures is made of the sum of two separate components: the expected net present value of economic benefits and costs which are associated to any actual or possible use of the research infrastructure (respectively B_u and C_u) and the expected net present value of other benefits referred to the social value of research discovery *per se*, regardless its possible use. The expected present value of non-use benefits (B_n) represents the pure value of discovery as a *pure public good*. The non-use term of Equation (1) can be expected to be always non-negative, as people are usually better off with discovery, or at least indifferent to it. Hence, our accounting convention is that the RI costs support all the use-benefits, while the pure value of discovery is an externality generated at no additional cost. This is only to simplify the presentation, and other accounting conventions can be used. The difference between use and non-use benefits is gathered from the field of environmental economics and we discuss more on this in Section 3.6.

Elaborating on equation (1), we argue that, apart from the pure value of discovery B_n, five main kinds of other measurable benefits stem from any RI project: the present values of knowledge outputs (typically related to scientific publications) S, technological spillovers T and human capital accumulation H. In many cases, particularly when dealing with large infrastructure projects, a fourth benefit, C, is related to the wider cultural effects of the project outreach activities. Additionally, as we move from fundamental to applied research and technological development, the service provided by the infrastructure could produce specific benefits to other users, A: effects on the health of patients of medical research facilities (e.g. accelerators used for hadrontherapy), environmental protection services (observatories studying natural hazards and technogenic risks), energy efficiency (centres carrying out research on renewable and efficient sources for the energy sector), testing of materials for private companies and license deals, etc. The present value of use-benefits can be written as the sum

---

[26] However, in Section 3.1 we argue that the salary of scientists, valued at shadow prices, can be used in some cases to value the benefit of knowledge creation, i.e. the value of publications and citations.

[27] See Section 4.

of all these components valued at prices (named 'shadow prices', see Annex 1) which reflect the social marginal value of a change of an output in the economy, i.e. the opportunity cost to the society of producing or consuming more or less of any good:

$$B_u = S + T + H + C + A. \tag{2}$$

Use-costs relate to the present value of capital K, labour cost L, other operating costs O, such as materials, energy, communication, maintenance, etc., and negative externalities E, like air pollution or noise during construction and operations. It is useful to decompose the labour cost term into the sum of the cost of scientific personnel L_s and labour costs of other administrative and technical staff working at the RI L_o:

$$L = L_s + L_o. \tag{3}$$

Thus, we have that the present value of measureable use-costs is:

$$C_u = K + L_s + L_o + O + E. \tag{4}$$

All terms of equations (2), (3) and (4) are expressed at shadow prices, and use-benefits and costs are properly discounted at the social discount rate r (see Annex 1). Therefore, the complete expression for the CBA model of RIs, over a time horizon spanning from 0 to $\mathcal{T}$, in simplified notation is:

$$NPV_{RI} = [S + T + H + C + A] - [K + L_s + L_o + O + E] + B_n. \tag{5}$$

Suppose for a while that NPV_n is completely unknown because we do not know anything about the pure value of discovery. The CBA test could produce three possible results:

- The net present use-value of the research infrastructure NPV_u is greater than zero, i.e. $PV_{B_u} > PV_{C_u}$, net of the unknown non-use good. It can also be written that $NPV_u > 0$;
- The net present use-value of the research infrastructure is equal to zero net of the unknown non-use effects, $NPV_u = 0$;
- The net present use-value of the research infrastructure is negative net of the unknown non-use effects, $NPV_u < 0$.

In the first two cases the RI passes the ex-ante CBA test if the evaluator *guesses* that the uncertain B_n would be at least nil, so that the total NPV_{RI} cannot be expected negative (within a range of associated probabilities). In other words, when the use-benefits of the RI are at least equal to the costs of producing them, in principle there is no further need to try to estimate B_n, as long as it can be excluded that non-use effects are non-negative. This is clearly a considerable computational advantage. The pure public good of discovery, if any, is still an unmeasured externality of the project, but the society gains or at least does not lose by having the RI. We suggest that for most RI in applied research and technological development the CBA test should be passed on these grounds.

In the third case the RI project passes the CBA test if and only if B_n is positive and large enough to compensate for the negative use-effects (costs and negative externalities). In this situation, we can no more avoid an estimation of the pure value of discovery. As mentioned, what is needed here is a *guess*, a conjecture of the possible impact of the discovery on social welfare. We discuss the nature of such guess below in Section 3.6 and we claim that empirical measurement of the social willingness to pay for the pure value of discovery in principle is possible, albeit with due caution, when the notions of existence value and quasi-option value, as established in other CBA fields, are duly adapted and borrowed in our context.

Once having defined the taxonomy of benefits of RIs, for each benefit there are two crucial steps to take in order to implement a social CBA. The first is quantifying the benefit itself, which has to be expressed in the form of a good for which there is a demand. If nobody is willing to pay anything for a good, literally that is not an economic good. The second step is valuing this good through a shadow price, which expresses the social value of a marginal change in the availability of the good.

The estimation of shadow prices is the main conceptual difficulty involved with the calculation of the NPV. Drèze and Stern (1987, 1990) prove that, in some cases, the shadow price of a good can coincide with its long run marginal cost of production, i.e. the social cost of increasing the production of that good by one additional unit, holding the production level of all other goods constant. An alternative approach to shadow price estimation is to consider the willingness to pay (stated by the project users or indirectly revealed through specific techniques), which entails the estimation of the social value of the good by summing the maximum amount people would be willing to pay to obtain that good. The latter approach is particularly appropriate to determine a monetary value for non-market goods. In some circumstance, the marginal social value of a good can also be obtained by a combination of the long run marginal cost and willingness to pay.[28]

The rationale for using shadow prices in place of observed market prices when evaluating the welfare impact of infrastructure projects relies on the fact that shadow prices better reflect the social marginal value of goods in an economy where markets are not perfectly competitive and efficient and market prices are likely to be distorted. In the case of research projects, the cost side is deeply influenced by the fact that the users of the service are the scientists, and a share of them are also the producers of the same service (this element is discussed more extensively in Section 3.1). As they are the direct users, they are project beneficiaries, thus they have an implicit willingness to pay for the project. But because they are also the service producers, they often pay for it in a non-monetary form, by contributing time and effort to the operation of the infrastructure. Also, generally investors do not seek in this context a monetary return for their stake in the project. Occasionally, one can observe situations where the owner of the infrastructure is paid by users, and the hidden economic values of a RI can be more easily revealed. Another reason why market prices are unlikely to represent a relevant signal for the decision makers of research infrastructure projects, is because the most relevant goods produced by the RI are either public goods, like non-excludable and non-rival knowledge, whose market prices typically do not reflect the opportunity cost of the good, or externalities, like technological learning, for which prices do not even exist. In general, this peculiar exchange of capital, labour, consumption and non-market goods, that we shall discuss below, is such that the price system does not work efficiently, and this is our case for using social CBA at shadow prices in this context.

In the next section we turn to a more analytical discussion of quantification and valuation issues regarding each of the above mentioned RI benefits, starting by the use-benefits and turning then to the pure non-use value of discovery. A number of research questions can be raised in this regard and are going to be addressed in what follows. As far as the RI knowledge output benefit is concerned:

- How can we measure knowledge outputs?
- How can we empirically estimate a shadow price for knowledge outputs?

These question can be answered by identifying empirical observable objects as knowledge outputs (i.e. scientific publications) for which there is a demand and valuing them from either a willingness to pay or long run marginal cost perspective (Section 3.1).

Research questions related to side-effects of research infrastructures are partially similar to the previous ones. For technological progress (Section 3.2):

- Is there a way to fully identify and measure technology progress associated to a RI?
- To what extent technology spillovers are actually an economic, rather than ‘pecuniary’ externality and how can they be valued?

As for human capital accumulation (Section 3.3), the research questions are:

- How can we measure the increase of human capital related to the participation of scientists and researchers to RI projects?

---

[28] Florio (2014a) discusses in detail the empirical issues involved.

- What is the shadow price of such increments in the stock of human capital?

The research questions related to the cultural effects on general population (Section 3.4), are:

- What is the outreach size of the research infrastructure?
- Is it possible to attribute a sensible social value to the wider cultural effects?

As for the services provided by applied RIs, the associated questions are:

- What is the demand for such services?
- What is the marginal WTP for them?

Finally:

- What is the social impact generated by a discovery itself, beyond its measurable scientific outputs, i.e. beyond the value of scientific literature and other observable use-effects ?
- How to value an intrinsically non-use effect?

The latter are the most difficult questions to be answered and to do so we will take advantage of previous theorization and experience mainly belonging to the literature of cultural and environmental economics. This issue is discussed in Section 3.6.

3 Evaluating the social benefits

3.1 Social demand and value of knowledge outputs

According to the mathematician and physicist Henry Poincaré (1908):

> ‘The scientist does not study nature because it is useful to do so. He studies it because he takes pleasure in it, and he takes pleasure in it because it is beautiful. If nature were not beautiful it would not be worth nothing, and life would not be worth living’.[29]

In the perspective of welfare economics, this amounts to say that there are people who have a preference for knowledge *per se*, just as there are people who like arts, or beautiful landscapes, or sports.

Answering the first of our research question – the measurement of knowledge outputs – may appear an impossible task. Knowledge *per se* is an intangible good and has a number of special features. First of all, there is widespread non rivalry. The fact that one knows a discovered fact does not subtract anyone else from potentially using the same knowledge. In other terms, the benefits derived from knowledge may extend to mankind in general. Second, there may be, to a certain extent, non-excludability, as some knowledge could not be patented or otherwise protected. Thus knowledge created by RI projects is often a public good and this creates a market failure.[30]

A second, and perhaps even more fundamental, market failure, is that knowledge creation is by definition characterized by the fact that ex-ante information is imperfect, as literally users do not know what they ‘buy’ when they embark in studying something unknown:

> ‘[…] the journey that the LHC has begun is an odyssey towards stranger spaces in which no one can predict exactly what will be met or where we will arrive. It is a search for unknown worlds which is carried out with complex cutting-edge technologies and guided by theoretical speculations […]’ (Giudice, 2010, page 3).

---

[29] Reported by Giudice (2010: 243).

[30] In economics, this implies that the quantity of public goods demanded by consumers does not equate the quantity supplied by suppliers. This creates a case for public intervention.

Some confusion may arise in this context from the lack of discrimination between the measurable output of the knowledge creation process, and the social value of discovery. The latter is a pure public good, the former is in principle a private good. We focus here on this private good, and delay to Section 3.6 the discussion of the marginal social value of discovery *per se*.

While new information generated at the RI is initially stored in computer memories[31] or in other technical supports, and obviously in the brains of the scientists, then it spawns a stream of specialized literature. The first wave of this literature may take the form of internal technical reports, preprints, eventually research papers in scientific journals and research monographs produced by all those scientists who directly use the RI and are involved in its operation and in the interpretation of first hand evidence. But besides 'insider' scientists, there are also 'outsiders' who are the rest of scientific community, including those working in other fields, who use the evidence provided by the experiment, as explained and discussed in the insiders' papers, to produce other knowledge. Even if not direct users of the RI, to some extent they are also beneficiaries of the project. Other waves of knowledge production can follow, with different scientists using the findings of 'second round' papers as a basis to write their own paper, and so on. This process is virtually infinite. In the next centuries there will be scientific publications which are the descendants of a genealogy started with the RI-linked literature, which is no more cited directly.

As suggested by other studies (e.g. Pinski and Narin, 1976; Martin, 1996) and CBA guidelines (JASPERS, 2013; European Commission, forthcoming) we accept the view that one empirical measure of research output is, albeit very imperfectly, given by publications. Presentations of results and findings at conferences are another form of communication used by scientists that should be added to this.

Bibliometric techniques, analysing the patterns of the scientific literature generated over time around a research infrastructure or its experiments, e.g. through keywords, citations, and other pointers, can be conveniently exploited to associate a measure of scientific output to the RI.[32] In practice, tracking knowledge output resolves in quantifying the knowledge outputs generated by the RI scientists (taken as level 0), papers written by other scientists and citing those of the insiders (level 1), other papers citing level 1-papers, and so on. A good understanding of regularities of such process is also the key factor in forecasting and ex-ante simulation. For an in-depth study of knowledge propagation in high energy physics, reference can be made to Carrazza *et al.* (2014), who analyse the citation distribution of papers related to different high energy physics infrastructures over a wide time span.

The average number of citations that a paper written by a scientist working at the research facility gets depends on many factors, such as the scientist's track-record, the scientific management and overall strategy of research, the competitive attitude and commitment of scientists, but also the luck in choosing the 'right' experiments, technical specificities of the research machines which influence the probability of discovery. Also, the number of citations tends to vary from one research field to another. Data on citations of journal articles indexed by Thomson Reuters in the Essential Science Indicators database from 2000 to 2010 show that on average each paper in the mathematics field receives around 3 citations, while papers in molecular biology up to 25 (Table 2).

---

[31] At the LHC something like a million Gigabytes per second of information is produced, 'sufficient to saturate every hard disk of the planet in about a day' (Giudice, 2010: 135). In life science there are databases for the human genome. Other examples pertain to chemistry and other disciplines.

[32] These techniques are discussed in a companion paper by Carrazza, Ferrara and Salini (2014).

Table 2 *Citation averages by scientific field*

Years	Average 2000-2010
All fields	*10.81*
Molecular biology	25.62
Immunology	21.81
Neuroscience	19.47
Biochemistry	17.25
Microbiology	15.79
Space science	14.30
Clinical medicine	12.93
Pharmacology	12.20
Environment/ecology	11.35
Psychiatry/psychology	11.26
Chemistry	11.19
Geosciences	9.70
Physics	8.97
Plant/animal science	7.74
Agricultural sciences	7.05
Materials science	7.03
Economics/business	6.22
Engineering	4.76
Social sciences, other	4.67
Computer science	3.75
Mathematics	3.48

Source: Times Higher Education 2011, http://www.timeshighereducation.co.uk/415643.article

For ex-ante projections, which is not our main concern in this paper, one could adopt empirical curves describing the dynamics of knowledge (identified as $S(t)$ in our CBA model) captured by publications and citations. An example includes a logistic function, leading to a differential equation of the Bernoulli form:

$$\frac{dS(t)}{dt} = \alpha \cdot S(t) \cdot \left[1 - \frac{S(t)}{\beta}\right] \tag{6}$$

where $\alpha > 0$ is an instantaneous growth rate parameter and $\beta > 0$ the equilibrium limit size of knowledge growth. There are several version of this simple non-linear differential 'epidemic' equation e.g. in the literature on innovation or mathematical biology. It has the well-known feature that the growth process is initially exponential and then slows down and asymptotically reaches a steady state (S-shaped process). Perhaps one could argue instead that the steady state will never be reached and direct citations will continue forever. Or one could suppose that, since knowledge is subject to obsolescence, after some time of stabilization there will be a decline in citations.[33]

For our purposes, preprints, publications and also contributions to conferences and any other product of knowledge produced by RI's scientists should be measured. It is important here to carefully distinguish between knowledge outputs and knowledge dissemination. When something is known, it can be transmitted in some forms: textbooks, articles in the press, documentary films, wikipedia entries, etc. The measurement of such dissemination activity in principle can be done, in terms of number of products, and its impact in term of readers, downloads, etc., or users in general. There is definitively a market for these scientific dissemination products, and while it may be not a competitive market, still it is a market. This is not however the right type of measure for knowledge creation: dissemination to a larger public is a parallel or subsequent process, partly dealt with in Section 3.4 on wide cultural effects, but the focus here is on the previous stage, when a flow of research output is

---

[33] These issues are discussed by Carazza *et al.* (2014) who discuss other functional forms.

published. A good understanding of regularities of such process is also the key factor in forecasting and ex-ante simulation.

We turn now to the second research question related to the valuation of knowledge outputs. As mentioned, a good has an economic value if somebody's welfare increases when its availability increases. This implies that the good is demanded by agents. What is very special in science is that the demand for the knowledge output production function of a RI project is driven by scientists in a given field who are often at the same time users and producers of knowledge. This does not happen for many other infrastructure services. Passengers of high speed rail demand the transport service, but are in no way involved in its production. Users of electricity have no stake in the nuclear power plants construction and operation. It is possible that in a cooperative firm, e.g. a rural electricity distribution company, customers are also the promoters and owners of the infrastructure, but in principle they do not build and operate it.[34]

The situation with RI projects is radically different: it is the scientific community that demands the project service, to prove or disprove theories through experiments, or to advance the knowledge frontier. The same people are also involved in design, construction, operation, interpretation of the evidence, elaboration on it, scientific discussion, until the new knowledge becomes accepted truth. Governments and the general public at large usually do not understand what exactly is at stake, and will be convinced to fund the costly RI usually based on the scientists' reputation and merit.

From this discussion, we maintain that one possible avenue to answer the question about the valuation of the benefit of knowledge output is to estimate the willingness to pay of scientists for the RI project. This may look as not very promising, as scientists in fact might not pay anything to use a research facility (e.g. in case of open-access centres) and are often paid instead to work there. Also, the research publications they write, the presentations they make at conferences or other research activities are produced for free and no or very limited earning is obtained from each of them.

The fact, however, that scientists are also the producers of knowledge, offers a different way to think to the value of this output. Most scientists are paid fixed salaries, and are relatively independent in the allocation of their time. Thus, when they spend some time on a research project, they have an opportunity cost, which is the fact that they do not work to an alternative project. If this opportunity cost is assumed equal to the average scientist's hourly compensation, then a reasonable proxy of the value of scientific output is its marginal production cost. This would be the time spent by scientists to make research and produce a paper, a preprint or other knowledge outputs, valued at appropriate shadow wages.[35]

If investigating the different strands of the scientific literature related to a certain RI can be feasible ex-post using proper bibliometric techniques, forecasting these benefits from an ex-ante perspective is more difficult. This is basically an empirical question, of high relevance, and for which the current experience is limited, but not such that it is entirely out of reach. If scientometrics works backwards, in principle it may work the other way round. In particular, the benefit transfer approach could be adopted. An analysis of the literature studying data of citations across time lags, controlling for heterogeneity across countries and scientific fields, could provide indications about the average number of publications and citations for a scientific experiment that could be expected for other experiments in the same field, and the temporal dynamics of such a knowledge production.

---

[34] There are some exceptions, that are to be found in some Non-Governmental Organisations and in local communities, or with distributed technologies as solar energy power. These exceptions are generally restricted to small scale infrastructures.

[35] We mention here a shadow wage because market wages are likely to be distorted: many scientists work hard if the project has an interest to them, and sink in the process much more time than it would be 'contractually' established. Conversely, it could also be the case of scientists with a low work productivity which are paid a salary higher than their opportunity cost. In fact, what is productive in the field of research is *per se* a controversial issue.

For example, as far as the average number of paper produced by a scientist is concerned, Newman (2004) examines data from computer databases of scientific pre-prints and published papers in physics, biomedical research and computer science. Newman shows that the average number of papers per author is between 2.55 and 6.4 over the five-year period 1995-1999, with the exception of the high-energy physics papers contained in the SPIRES database, in which the figure is higher and around 11.6 papers per author. This is explained by the higher number of authors co-authoring papers in this scientific domain (an average of 9 against 2-3 for other subject areas), but also by possible double counting of preprints and publications by the same author. These data suggest that each year an average of 2.32 preprints and papers are produced by high-energy physicists. The considered databases for other sciences include only publications or preprints, and thus the number of papers per year is lower: 1.28 by biomedical scientists and less than one by researchers of other fields. Even if these data are not comparable among themselves, they give an idea of what type of statistics can be obtained with scientometrics and then used to estimate the marginal cost of knowledge outputs.

Table 3 Statistics on selected databases of papers – 1995-1999

	Biomedical research[36] (publications only)	Astrophysics (preprints only)[37]	Condensed matter physics (preprints only)[38]	High Energy Physics (preprints only)[39]	High Energy Physics (preprints and publications)[40]	Computer Science[41] (preprints only)
Average authors per paper	3.76	3.35	2.66	1.99	8.96	2.22
Average papers per author over the 1995-1999 period	6.4	4.8	3.65	4.8	11.6	2.55
Yearly average papers per author	1.28	0.96	0.73	0.96	2.32	0.51

Source: Authors’ elaboration on Newman (2004)

The marginal production cost of a paper would capture only part of the total value of knowledge output. In fact, the value of knowledge is made of two components: the social value of producing new information *per se* plus the social value attributed to the degree of influence of that piece of knowledge on the scientific community. If the former is captured by the number of papers written and valued through the marginal production cost, the latter is reflected in the number of people would read the paper (reflected e.g. by downloads from an electronic repository) and eventually the number of citations a paper gets. Using citations as a measure of the significance of a scientific paper is an imperfect but widely accepted approach and we accept the view that – on average – citations reflect the social recognition and esteem that the scientific community acknowledges to the paper (Hagström, 1965; de Solla Price, 1970). It is therefore reasonable attributing a statistically higher social value to a paper that has received a higher number of citations in a given period compared to a less cited paper in the same field.[42]

A shadow price of citations need to be estimated, and by analogy with the value of paper production, this could be the opportunity cost of time employed by a scientists to read and understand someone else’s paper and decide whether to cite it or not. This time can vary from few minutes to many hours

---

[36] Database used: Medline.

[37] Database used: Physics E-print Archive.

[38] Database used: Physics E-print Archive.

[39] Database used: Physics E-print Archive.

[40] Database used: SPIRES.

[41] Database used: NCSTRI.

[42] This has nothing to do with the evaluation of the intrinsic quality of the paper, which is not a statistical measure.

or days, depending on the type of paper, its length, topic, the experience of the citing scientist and other variables.

Thus, going back to the CBA model shown in equation (5), on the one side the total cost of the RI is given by the sum of investment and personnel (scientific and other labour costs), other operating costs and possible negative externalities, while the primary benefit of RIs is given by scientific outputs (publications and citations) cumulatively generated by the project, valued with the time of scientists used to produce them and a shadow wage. This may seem a peculiar way to represent the cost-benefit balance of the project, because we have some costs that are *de facto* wiped away by benefits: the salaries paid to scientists working at the RI are a cost for the infrastructure, but the cost of their time employed for doing research and producing knowledge is a benefit for the society. This view is however consistent with the standard CBA assumption that in some contexts marginal costs are the best proxy of marginal benefits.

To sum-up: first, we need to consider time allocated and shadow wages not only of scientists directly working at the research infrastructures, but also those who use and elaborate on it to produce new knowledge, so as to capture the cumulative process of knowledge output production; second, the value of citations received should be added to the value of paper produced in the first, second, … *n*-th wave. In this way, extremely valuable, seminal papers, that introduce brilliant new ideas that receive a high number of citations are valued more than obscure non influential papers.

Even if citations are an accepted way to detect impact, we clearly cannot go as far to say that all the *n*-waves of citing papers would not have existed without the initial RI scientists' papers. In other terms, the shadow price of the generated knowledge output cannot be simply multiplied for the number of papers in the genealogy, since it would be an exaggeration to state that there is a one-by-one relation between knowledge units produced in the first round, and those produced subsequently. An easy, but effective shortcut to deal with this issue is simply to divide the value of papers produced by outside scientists by the number of references contained in the same papers, as if each contributed in the same way to the new knowledge output.[43]

Thus, after the mutual cancellation of first round costs and benefits, the benefit of measurable knowledge output is the value of citations that papers of RI's scientists receive, plus the total value of paper production and citations of the subsequent waves of papers. As for our interpretation, a crucial issue to value knowledge output is to correctly and fully identify and value the citations and the output of outsider scientists deriving from the output of the insiders. In this way, the net present use-value of the RI project would be greater than zero only if the investment and operation costs are less than the value of citations and indirect knowledge output, plus other use-benefits accumulated in the long run and mentioned in equation (5).

In Annex 2 we provide the details of a simple model to value the knowledge outputs associated to a RI in a CBA framework based on these ideas. Some preliminary empirical tests show that even for major scientific enterprises it is unlikely that the *S*-benefit takes a very high value relative to other benefits and costs in equation (5), but it is important to understand that the main social benefit here lies not in this value, but in the pay-back of scientific personnel cost through the valuation on the knowledge output they produce, and this is a major contribution in terms of CBA of the research infrastructure.

---

[43] Abt and Garfield (2002) analysed 41 research journals and found that the number of references is generally between 20 and 70 for biochemistry and molecular biology, between 20 and 50 for physical sciences and between 5 and 60 in medicine. Also, they noted that there is a significant linear relationship between the average number of references and the normalised paper lengths.

3.2 Technological externalities

Earlier literature on research infrastructures often reports that their impact on technological progress is a social benefit to be considered as part of the broad picture of growth and welfare effects, additional to the advancement of fundamental knowledge. Building a new large and complex infrastructure or carrying out an experiment at the scientific and technological frontier can be an important source of innovation (Lederman, 1984; Kay and Llewellyn Smith, 1986; Mansfield, 1991; Technopolis Group, 2013; Del Bo, 2014).

A well-known example of technological spillover is the invention of the World Wide Web at CERN in 1989, initially conceived as a means to improve the sharing of information between scientists working on CERN experiments. Substantial progress is now expected with other software developed at CERN including the worldwide LHC computing grid project, which allows to link, distribute and analyse a volume of information that currently exceeds the capacity of any computing facilities in a single site (Boisot, Hoffmann and Nordberg, 2011). Grid computing is widely used in business and science context, specifically for science applications that require large data processing capabilities, such as climatology, astronomy, biology and others (Giudice, 2010: 138) and has attracted the attention of the core players in the computing industry. There are similar narratives about the technological jumps related to other research infrastructures. For example the Global Positioning System (GPS) was originally intended by the US Department of Defence for military applications before being made available for civilian use in the Eighties;[44] a wide range of new materials and tools stem from space technologies needed for the NASA projects, such as the 'memory foam' able to deform and absorb pressure and to return to its original shape: invented to improve the safety of aircraft cushions, it is nowadays used for helmets, mattresses or wheelchair seats.[45] Other technologies originated at the European Space Agency, the European Southern Observatory, European Synchrotron Radiation Facility, European Molecular Biology Laboratory, etc.

Technological spillovers might occur from the work carried out not only by the RI staff, but also within the firms and laboratories along the RI's supply chain. Firms often do not have ready-made solutions to the types of problems that arise when involved in the design, construction or operation of a complex, high tech scientific instrument, for example related to the need of increasing precision of mechanical components, weight, other physical properties of materials, design of electronics, etc. When a procurement contract for the RI is signed, an intense collaboration process between the supplier and the RI itself gets started aimed at effectively designing, testing and manufacturing the required product or service. These efforts give firms the opportunity of learning something new and to use the new skills for producing further technological advancement to be exploited.

The analytical issue involved in estimating the technological impact of RIs is two facets, as mentioned in Section 2: i) how to identify and measure technological progress, and ii) how to value it. As a first step it should be ascertained whether the R&D costs involved in producing an innovation have been paid entirely by the client or by the firm itself and not directly recouped through the main procurement contract. If the research and development cost is fully internalized by the firm, and is then repaid by the procurement contract, there is no identifiable 'first round' externality and the services provided by external firms will appear as a cost rather than a benefit in the analysis. However, this does not bar 'second round' effects to occur. The learning-by-doing process triggered by attempting to solving a practical problem could generate further innovation, that would have been impossible without the initial pull effect. Innovation spilling over the scope of the initial procurement contract can be, at least to some extent, attributed to the knowledge acquired on the job.[46]

---

[44] http://geography.about.com/od/geographictechnology/a/gps.htm

[45] http://science.howstuffworks.com/innovation/inventions/top-5-nasa-inventions.htm#page=1

[46] In this vein, Giudice (2010: 109) reported that: 'Many of the companies that worked for the LHC project are now using the new skills learned in the process. For instance, one company is producing superconducting material for medical magnetic resonance imaging and another has applied a special production process started for the LHC to manufacturing automobile parts'. The UK Science and Technology Facilities Council (2010) mentions other examples of companies who have benefitted from expertise gained by working and interacting with the RI (in that case the Darensbury Synchrotron Radiation

Learning-by-doing as an externality of R&D activities has received great attention in the economic literature on endogenous growth. Arrow (1962) used the concept of on-the-job learning as determinant of technological change to investigate the economic implications of knowledge growing in time. In this model, the rate of growth of technology depends on the rate of growth of capital, reflecting the fact that productivity increases when the cumulative production grows, thanks to the learning-by-doing effect. Higher productivity and, hence technological change, are expected to lead to higher profits. Of course, the absorptive capacity of the firm, this being its ability to recognise the value of new information or skill and to assimilate it and apply it to increase its profits, is a critical factor (Cohen and Levinthal, 1990).

There exists a vast literature analysing the relationship between academic research and industrial innovation activity. Just to mention two studies, an econometric analysis by Jaffe (1989) found a significant positive impact of the university R&D on industrial patenting in 29 US states, see also Bacchiocchi and Montobbio (2009). On the same line of thinking, Cowan and Zinovyeva (2013) have recently analysed the effects produced by the opening of new universities in Italy during 1985 and 2000 on regional innovation, in terms of the number of patents filed by firms, and confirmed the existence of a positive relation. Other studies show that university research also positively affects firms' product and process innovation (Acs *et al.*, 1992; Feldman and Florida, 1994).

The empirical literature focusing on the technological spillovers of research infrastructures is less extended. The first studies were drafted in the Seventies by the NASA in the US and CERN in Europe. These studies usually rely on a qualitative methodology of analysis and case studies, developed through desk research, in-depth interviews and surveys. One of the most cited studies was carried out by Autio, Bianchi-Streit and Hameri (2003). The authors investigated the learning benefits gained by European firms that had participated in CERN's procurement activity between 1997 and 2001. A sample of firms was selected from the total number of suppliers to CERN during the considered period (6,806 firms), excluding those companies whose total order did not exceed CHF 25,000 and which provided only off-the-shelf products or very simple services.[47] The sampling process led to 612 companies that were assumed to have supplied a noticeable technological development or innovation components.[48]A survey was then submitted to this sample of firms and based on the answers provided by the respondent firms (154), the authors found that the benefits associated with procurement activity can be in terms of significant technological and market learning (respectively 44% and 36% of firms), international exposure increased (43%), new products developed (38%), new markets opened (17%), new business units established (14%) and new R&D units started (13%). Respondents also declared that, without CERN, they would have had poorer sales and technological performance (52% and 41%), poorer performance in valuation growth (26%), and lower employment growth (21%). More recent data refer to the ATLAS experiment, see Autio *et al* 2011.

An additional innovation outcome that might be produced by RIs is the creation of spinoffs, aimed at commercialising the facility's research breakthroughs. NASA has been tracking its spinoffs since 1976 and has now a database including 1,800 spinoff case studies,[49] the majority of which associated to the Langley Research Centre, the Johnson Space Centre and the Marshall Space Flight Centre. In general, an average of 48 spinoffs are generated every year by NASA research infrastructures.

If it is largely accepted that building a complex and often unique device for research infrastructures operating at the very edge of the science, is very likely to generate widespread technological and

---

Source), such as a company that improved its own precision magnet measuring equipment allowing it to improve its competitiveness.

[47] Like catering or training.

[48] They represent less than 10% of the total number of suppliers to CERN, but 56% of CERN's total procurements during the period (CHF 1,197 million). Their fields of activity spanned from information technologies, to electronics and material sciences (related e.g. to magnets, cryogenics and vacuum technologies).

[49] NASA online spinoff database, available at http://spinoff.nasa.gov/spinoff/database.

industrial innovation, detecting and quantifying all technological externalities spilling over the RI is particularly challenging. From an ex-post perspective, only systematic surveys and detailed case histories through interviews in the supply-chain firms can help establish their existence.

From an ex-ante point of view, one reasonable way to forecast the possible size of technological spillovers of the RI under assessment is to take already existing similar RIs as a benchmark and rely, as far as possible, on the opinion and expectations of experts about the similarity or dissimilarity of technological patterns. The probability of error can be tested through a fully-fledged risk assessment (see below).

Having found histories and trajectories of economic innovations, valuing them is a separate endeavour. We agree with Boisot and Liyanage (2011: 245), who state that:

> 'Science and technology together create useful knowledge. If the stress in science is on the knowledge itself – knowledge creation for its own sake – the stress in technology is on its utility'.

In other words, while a value can be attributed to knowledge outputs *per se* (the social value of producing a paper), regardless of its actual influence, the value of technology is linked to its actual usability and to the concrete advantages it leads to. Ideally, one should look at the social profits generated by them, catered from the company's return on sales. Being j the number of companies benefitting from technological spillovers over time t, Π_{jt} their incremental shadow profits (i.e. profits at shadow prices) directly imputable to the spillover effect, and s_t the discount factor, the present value of technological externalities is expressed as:

$$T = \sum_{j=1}^{J} \sum_{t=0}^{\mathcal{T}} s_t \cdot \Pi_{jt}. \qquad (7)$$

The firm's return on sale reported in balance sheets can be often taken as a proxy of social profit in competitive markets; in distorted markets, so that observed prices do not reflect the real opportunity cost of resources, the profit has to be derived as the difference between the firms' total income or cash inflow and operating costs, all valued at shadow prices (see Section 2).

Our approach is broadly in line with the empirical literature, where R&D spillovers and externalities are captured through variations in the private profit margins (e.g. Mansfield *et al.*, 1977; Hall *et al.*, 1999; and Hall *et al.*, 2009), and it can be adopted also in a CBA framework subject to the important proviso that only variations in profits that are ascribable to the activities carried out by RI's supplier are considered. This is easy in case of new spin-off companies created to commercialise a technology associated with the RI, whose benefit is reflected in the cumulative profit made by the company during its entire lifecycle.

But the causal link between the activities carried out for the RI and future profits due to those specific activities might not be obvious. To solve this issue, the increase of profit in principle should be assessed against a counterfactual group of companies, operating in the same sector and sharing other characteristics with the companies that actually worked for the RI, in order to control for selection bias. The set of techniques typically used for implementing a counterfactual impact evaluation,[50] which are well established especially in the evaluation of the effects of government subsidies on private R&D in the European Union,[51] can be relevant also in the RI context.

While ex-post a survey to companies within and outside the supply chain of the RI could be set up, one crude form to value the technological progress ex-ante would be to use a 'benefit transfer' approach, i.e. giving a money value to innovation indicators related to a specific project plugging in a

---

[50] Difference-in-difference, discontinuity design, matching approach.

[51] See for example Gadd et al. (2009), Mouqué (2012) and ASVAPP (2012).

value estimated from existing knowledge elsewhere. This is imprecise, but better than using only guesses. For example, if a range of estimates about the marginal impact generated from R&D activities on firm profitability due to its direct exposure to the RI is available, we may transfer this information to the RI project and use it as a proxy for the social benefits (at shadow prices) of the project on the supply chain.

The idea of tracking patents linked to the development of RI projects,[52] as suggested in some literature (see among others Scherer, 1965; Schmookler, 1966; and Hall *et al.*, 1986), could provide a useful but only partial indication of the total innovation produced. As a matter of fact, not all innovation generated both by the research infrastructure owner, its scientists and technical staff, and by firms that supply the RI with innovative materials and technologies, is patentable or might be actually protected by a patent. Only in the (usually unrealistic) case that all innovation spilling over the R&D activity covered by the procurement contract is patented, the economic return generated from such patents would fully reflect the value of technological externalities. To simplify, one could transfer the marginal economic return of a generic patent in a country or region to the RI project.[53]

In general, the increase of profits ascribable to the RI against a realistic counterfactual should provide a most comprehensive measure of technological spillovers, accounting for the benefits related to the production of a new marketable product, the commercial exploitation of a patent, the increase of productivity, and also increase of visibility and corporate's image. As stated by SQW Consulting (2008: 31):

> 'For specialist firms supplying equipment that is installed in the [research] facility there [can be] marketing benefits both from visitors seeing the company's badges on pieces of equipment and from word of mouth discussions between scientists. Scientists from different institutes talk to one another and reputation from a successful installation in one facility can be a useful marketing tool'.

Even whenever the value of technological externalities is determined, the cumulative second, third and *n*-th round effects to other companies which have not been directly exposed to the RI but that experienced imitation benefits would not be captured. Imitation is in fact a powerful multiplier of any technological externality.

The approach here suggested to value technological externalities cannot be confused with the way sales or increased efficiency and performance are generated by procurement. Some studies define the economic benefit of technology transfer as the sum of the increase of turnover and saving in production cost generated by, but independent from, the procurement contracts. In the context of CERN, for example, Schmied (1979) and Bianchi-Streit *et al.* (1984) analysed the supply chain of CERN respectively in the periods 1955-1978 and 1973-1982. The former study, based on data collected through interviews to a sample of 134 European firms (127 respondents) suggests that the 'economic utility' ratio was in the range 1.4 and 4.2 with an average of 3. This figure would indicate that for every Euro spent by CERN in a high-tech contract, a company receives around 3 Euros in the form of increased turnover or cost savings. As stated by Schopper (2009: 150):

> 'this implies very crudely that in a laboratory such as CERN about one quarter of the budget is spent on high-tech products and consequently around three quarters of the overall public spending is eventually returned to industry'.

---

[52] Either through names of inventors associated to patents, or of firms, or keywords.

[53] The OECD study 'Turning science into business' (2003) shows how skewed the distribution of licensing revenue per license is: 'While some [Public Research Organisations] in the United States generate several million USD from licenses, the average value of each license in 2000 was USD 150,000' (OECD, 2003: 16). In Switzerland the average revenue per license is much lower, EUR 45,000. The EIB (2013) reports an average yearly revenue per licence in Germany of EUR 55,000, with a higher average (EUR 200,000) for licenses of the Max Planck Institute. This issue is also discussed by the new EC Guide to CBA (European Commission, forthcoming).

Improving the sampling and extrapolation methods, Bianchi-Streit *et al.* (*ibidem*) obtained the same average utility/sales ratio. Other studies (mentioned by Autio *et al.*, 2011) report ratios of total value added to contract value of 2.7 in the case of the European Space Agency, and between 1.2 and 1.6 for Big Science centres.

However, these calculations seem to implicitly assume that the value of the externality can be computed simply as increased sales and decreased costs. In general, however, we maintain that it is not change of sales that needs to be considered, but the change of *net* output (i.e. profit) at shadow prices. If shadow prices are simply estimated equal to market prices, this would be the net present value of the additional gross profit along the supply.

So far we have focused on technological externalities produced on the RI's supply chain, but this benefit can in principle arise also within the RI itself, as the above mentioned World Wide Web example shows. If the innovation produced is patented and a license for its exploitation is granted to another institute or firm, the technological progress is fully internalized in the RI's cash inflow. However, just like for firms, not all innovation may be patented and give rise to economic returns. Even when the RI is provided with dedicated staff in charge of detecting possible commercial opportunities linked to technologies developed inside the RI, often significant investment cost would be needed to bring the technology outside the RI and generate socio-economic benefits. Hence, it should not be surprising to discover that most of innovation produced in very large RIs for fundamental science remains hidden. Some preliminary empirical testing and interviews to insiders at CERN, ESFRI and the EC – Directorate-General for Research and Innovation suggest that the variability of the *T*-benefit is very high across fields of science, with a possibly a maximum for high energy physics and a minimum for life sciences. Further empirical research on this variability however is needed.

Since we argue that the values of technology depends on its economic utility, innovation that remains unexploited and does not produce an actual increase of profits cannot be valued as a technological use-benefit of the project. It can however be examined with reference to the pure value of discoveries (see Section 3.6).

3.3 Valuing human capital formation

A large-scale research infrastructure attracts junior scientists from several countries. However narrowly-defined, perhaps even mechanical and possibly unattractive, is the task they are asked to carry out, such as experiments' data collection and processing, the motivation of students and junior researchers in doing so lies in the willingness to be 'part of the show, to be a player associated with one of the world's biggest scientific experiments' (Boisot and Bressan, 2011: 206). Big research ventures are stimulating and dynamic places, giving to young fellows and Ph.D. students the opportunity to interact with other students and scientists from different contexts. Thanks to observation, imitation and practice on the work place, but also to the participation in meetings, seminars, conferences and other events, most of students coming to the RI for short periods benefit from the development of their skills, ranging from technical and scientific abilities, to personal ones, related for example to the improvement in the communication, managerial, negotiating and organisational capabilities. Similar human capital accumulation benefits can also be observed with smaller scale research centres and laboratories.

With some adjustment many skills acquired at the RI could find practical application even in domains outside their own research field. Schopper (2009) states that around 40% of students working at CERN eventually go to industry, even 60% according to Maiani (2012). Camporesi (2001) analysed the careers of more than 600 diploma, masters and Ph.D. students involved in one of the Delphi experiment carried out on the LEP accelerator between 1982 and 1999. While 57% of them continued doing research and teaching in the academic context, 43% found their first occupation in the private sector, especially in the field of high technology and computing. If the analysis is limited to Ph.D. students, the share of those who remain in universities and research institutes is 67%.

In contributing to the training of young scientists worldwide, in fact most RI projects are similar to research universities, with one main difference: usually students do not pay a fee for their on-the-job training: in fact, the opposite may be true, as students are often supported by a fellowship. There seems to be a clear externality here, and the shadow price is, as usual, the marginal social benefit of such on-the-job training.

In this perspective, several insights from the economics of education can be utilized to gauge the contribution of research infrastructures to the increase of human capital available to society. Both theoretical and empirical analysis[54] suggests that secondary and higher education and training positively contribute toward economic growth by increasing the productivity of the labour force. A quasi-experiment perspective would be needed to assess the effect produced by the RI on young scientists. In its more appropriate form, this would imply tracking careers of cohorts of students in the long run and matching data on careers of scientists involved in RI projects with those who have not been involved. There are several econometric issues to be addressed: endogeneity, if there are reasons to believe that the best students might have been given a priority in getting a post in the flagship Big Science projects, and causality. In other terms, the difficulty is how to be sure that the higher earnings possibly observed for students who have participated to a training programme at the RI are caused by their participation to such a programme rather than other factors. Furthermore, in many situations controlled experiments are not a feasible solution because data on careers for RI students and of a suitable control group are not easily available.

In the absence of quasi-experimental evidence, the standard approach would be to set up an econometric model so as to estimate the marginal effect of human capital formation on the earnings gained in the entire life time. One of the most used econometric model is based on the Mincer's human capital earning function (1974) which disaggregates individual earnings, expressed in a logarithmic form, into a function of an education term (as given for example by the number of years of education, or the degree) and experience (as measured by the number of years of work since completion of schooling), a constant parameter and an error term. Instrumental Variables (IV) are usually used to reduce the correlation between the explanatory independent variables and the error term and, thus, to better 'explain' the net contribution of the independent variables on the dependent variable. Instrumental variables could relate to the student's country of origin, gender, race, parents' level of education, quality of the education, and so on.[55]

A review of the literature carried out by Card (1999) shows that IV estimates of the return to education are in the range of 2.4%-11%. A European survey by Psacharopoulos (2009) reports a minimum private return to higher education[56] of 2.1% in Croatia to more than 20% in Czech Republic, Poland and Portugal (2004 data), with an average of 10.2% in 31 European countries.[57] It also finds that there is a weak inverse relationship between the returns and the country's per capita income. The evidence for the return to different higher education facilities is more limited. A study on UK faculties (O'Leary *et al.*, 2005) indicates high returns associated to maths and computing (21.1%), education (19.4%), medical related (17.4%) and engineering (15.8%) degrees; returns to education in sciences, business and economics and social sciences are around 12%; the lowest return is associated to arts (4.1%).

In addition to human capital formation for students, some skills can also be acquired by scientists, engineers and technical staff working at the RI. Just like firms of the supply chain, RI's own employees with responsibilities over the design, prototyping, assembly and manufacturing of products for guaranteeing or improving the RI's functioning can benefit from on-the-job learning. This is an

---

[54] A very selective sample includes Schultz (1961), Mincer (1974), Psacharopoulos and Patrinos (2004) and Blaug (1987).

[55] An alternative approach to instrumental variable regression models is to study education attainments and earning outcomes for siblings. This should allow to reduce or eliminate the bias caused by unobserved differences (Card, 1999).

[56] Private return is defined as the increased earning (after tax) for an individual that has achieved tertiary education net of what he/she has paid to attend the education institute, relative to the control group of people with secondary level of education. In other literature, returns to education are calculated in different and not consistent ways.

[57] The European Union Member States except Malta, plus Iceland, Norway, Switzerland and Turkey.

additional benefit produced by the RI-project, which is incorporated in the increase of productivity of the RI itself and reflected in an increase of labour costs, as long as the staff members keep working there. But if they leave the RI, the increase of earnings they would get in performing another job compared to what they would have get without their experience at the RI is a positive externality of the research project.

In practice and similarly to the valuation of human capital accumulation of RI's students, when implementing a CBA of the RI project, ex-post the technical employees who, at a given time, left their job should be identified; a quasi-experiment should then be carried out to determine the marginal increase of income they obtained from working in another place compared to a control group of people, having the same job and the most similar features as the former RI-employees, except for their previous experience at the RI. Alternatively, an econometric model with control variables could be used. Ex-ante, with no actual data at disposal, the project analyst can only rely on assumptions on the number of RI-staff who will decide to exploit their new skills elsewhere, and the marginal increase of earnings ascribable to the RI, to be tested in the risk analysis. In some cases, inference from other contexts may be helpful (again under a benefit transfer approach).

The present value of human capital accumulation benefits produced by the research infrastructure project can then be defined as the sum of the increasing earnings or income, I, gained by RI's students and former employees, commonly indexed by z, since the moment (at time φ) they leave the project.

$$H = \sum_{z=1}^{Z} \sum_{t=\varphi}^{T} s_t \cdot I_{zt}. \qquad (8)$$

In summary, it is worth to note that the lesser RI-specific skills acquired by students and technical staff are, the larger the human capital formation benefit could be (Boisot and Bressan, 2011).

Preliminary testing of our approach with the LHC case study shows that the *H*-benefit can be substantial compared to other ingredients of equation (5), mainly because of the high number of students and junior researchers attracted in a large RI, and above all because of the long duration of the effect.

3.4 Outreach and cultural impact

Many research infrastructures, particularly, but not only, large-scale ones, regularly conduct a programme of outreach events and services aimed at informing the public on advances in science and technology. The organisations or institutions that operate the RI, or external institutes/companies on their behalf, often work to make the RI's site the destination of 'science tourism', e.g. by setting up permanent or temporary exhibitions, organising guided tours, granting the access to special events, open days, lectures and workshops.

Even less known facilities attract visitors. For example, a report exploring the social and economic impact of the Daresbury Synchrotron Radiation Source (England) highlights the role played by this infrastructure in extending the public awareness of science, engineering and technology (Science and Technology Facilities Council, 2010). Since 1995 the Daresbury laboratory has committed an increasing volume of resources to public outreach activities at local and regional level. This produces every year a flow of about 3,000 visitors, and, additionally, 3,000 school students per year are involved in ad hoc programmes and activities either at the Laboratory or by the Laboratory's staff at schools.

The size of cultural impact associated to Big Science projects can be much larger. The US air force area of Cape Canaveral is probably the most popular RI for Big Science. The Kennedy Space Center (KSC) Visitor Complex offers a variety of attractions, like the Rocket Garden, a 3-D theatre, exhibits of artefacts and robots, a memorial dedicated to astronauts, visit to the Space Shuttle Atlantis, activities simulating the astronaut training and much more. As the official launch site for NASA, it also offers close viewing of rocket launches. The Complex is one of Florida's most popular tourist destinations and it hosts more than 1.5 million visitors[58] from all around the world at an admission fee of USD 43 for adults and USD 33 for children.[59]

There are standard CBA approaches to evaluate cultural tourism to museums or other recreational activities, like visiting a natural park. We suggest to exploit these methods for scientific tourism as well. These approaches usually rely on the estimation of the willingness to pay, W, of the general public by type of beneficiaries ($g = 1, \dots G$) for visiting the RI project. Hence, we can express the benefit of outreach activities as follows:

$$C = \sum_{g=1}^{G} \sum_{t=1}^{T} s_t \cdot W_{gt}. \qquad (9)$$

The travel cost method is a well-established approach for valuing the willingness to pay of people for a desirable good, in this case a visit to the RI.[60] It consists in evaluating a good through the full travel cost incurred in its consumption, including the cost of trips (fuel, train or airplane ticket, etc.), the opportunity cost of time spent in travelling, the cost of accommodation, food, souvenirs and so on. Given the number of visitors to the site in a given time period and the marginal economic cost of a trip, the demand curve can be derived and the willingness to pay for a visit estimated.

The travel cost method however is affected by a limitation that should be carefully dealt with. It has to do with the apportionment issue arising whenever it is reasonable to assume that a trip is made for different reasons (multi-purpose trip) and not for visiting a specific RI. Actually it could be arduous to disentangle the willingness to pay of visitors for a given infrastructure when more than one attraction are located in the same site or in the same area. The full travel cost of people going to Florida to visit, among other things, the KSC, should not be entirely imputed to the KSC. An apportionment assumption is then necessary to account for the RI-related cultural impact, so as to estimate as far as possible the relative contribution of the RI on the total flow of visitors.

It is occasionally mentioned (e.g. COST Office, 2010) that some economic opportunities are likely to arise around the tourism demand created by the existence of the RI. Commercial and accommodation activities and other business opportunities near the infrastructure could benefit from the higher flow of customers. These non-technology driven spillovers are what are generally called ‘pecuniary externalities’, i.e. externalities operating through price adjustments in goods, properties and land. It is usually difficult to find a direct causal relationship between a project (of any type) and prices adjustments, so that this kind of wider effects are generally not accounted for in a social CBA. Also due to the limited relevance of this effect for the majority of research infrastructure projects, our suggestion is usually not to value it.

Besides visits in person, participation to activities on social media, television audience and website visiting are further indicators of the size of the cultural impact produced by the RI, also to be included in term W_{gt} of equation (9). These can be quantified through proper techniques commonly used by marketing specialists, e.g., via the number of ‘tweets’ or followers in Twitter, posts or pages in Facebook, subscribers of the YouTube dedicated channel or number of views of a video, the estimated number of people watching an event on TV, number of blog conversations, analysis of the volume of web traffic, registrations on the RI official website and so on. While their incremental costs is easily quantifiable, the incremental cultural benefit associated to these behaviours may be immaterial. Prices either do not exist, as most online services are provided for free, or do not fully capture the social economic value of this sort of ‘virtual tourism’.

Revealing the tacit willingness to pay for social network sites has been receiving increasing attention in the literature and the number of studies exploring the factors that determine the users’ willingness to

---

[58] http://media.kennedyspacecenter.com/kennedy/quick+facts/

[59] Tax excluded (http://media.kennedyspacecenter.com/kennedy/quick+facts/).

[60] See Florio (2014a) for a review of methods to estimate the willingness to pay for a good.

pay is growing. In analysing the structure of social networks, Westland (2010) stated that when a network reaches a certain critical mass a willingness to pay for network membership arises. Han and Windsor (2011) found that the trust generated from social activities favourably affects trust in business transactions on social network sites, thereby influencing users' willingness to pay. Vock *et al.* (2013) modelled the willingness of social networks' users of paying a premium fee for benefitting of upgraded services, compared to regular membership for free, and found that social capital and the perception of people as being bonded together in a coherent unit[61] result in specific values for members, which in turn positively affect their willingness to pay.

Contingent valuation is a method which is widely used to attach a monetary value to non-market goods. One of its versions consists in asking people to state the maximum amount of money they would be willing to pay for obtaining a good, or to accept as compensation to give away a good, contingent to a given scenario. However, empirical studies show that when consumers are accustomed to receiving an online service or content for free, their willingness to pay is very low or nil. A survey carried out in 2002 concluded that, when asked to pay for access to a site, only 12 percent of US Internet users indicated they would pay, while 50 percent would find a free alternative, and 36 percent would simply stop getting the service online (Crosbie, 2002). These results were confirmed by other empirical studies, such as a survey to more than 800 Hong Kong residents, which revealed that only few people were willing to pay for online new content (Chyi, 2005), thus explaining the failure of many news sites that charge a subscription fee; the willingness to pay is on average positive only for younger users and those spending more time reading newspapers, while, interestingly, it seems not to be affected by income levels.

Difficulties in obtaining values of willingness to pay through contingent valuation have been experienced in the cultural sector too (Snowball, 2008). In this context, the choice experiment or conjoint analysis methods are considered more useful that traditional contingent valuations. While based on stated preferences, these techniques imply asking a sample of population to choose or rank different combinations of attributes of the same good (a museum, an archeological site, etc.), where price is included as an attribute. This enables a more effective assessment of preferences in terms of willingness to pay both for each attribute and for the whole good. The same techniques could be usefully exploited also to our purpose, in order to attempt to value the public interest for the RI.

Finally, a further wider impact of large-scale RIs should be discussed: namely the demonstration effect of research infrastructures in attracting brilliant young students to scientific education. The PUSET programme of the aforementioned Synchrotron Radiation Source project explicitly aims at contributing to the formation of a knowledge-based economy, by encouraging interest in science and increasing the uptake of physical sciences and maths at school level. Young students, particularly the most clever in their generation, and less constrained by income of their families, have different opportunities of education choices. They may choose the humanities, law, finance, engineering, biology, or physics, for example. These choices are not purely random, and they are influenced by several factors. If the Big Science projects shift the perceptions of opportunities and of choices, is this a social benefit that can be valued? This is clearly a difficult question to answer. It is reasonable to believe that choosing a career in science is not *per se* more valuable than a career in economics, philosophy or arts. An incremental benefit can be acknowledged only if a lower level of education would have been attained in the without-the-project situation, regardless of the field of activity. As a wider and hardly quantifiable effect, its socio-economic valuation can be excluded in principle from the CBA, but it is nevertheless worth mentioning.

Preliminary empirical testing of these ideas shows that the *C*-benefit can go from relatively significant for very large RI such as the CERN to virtually negligible for health research facilities such as CNAO (Pavia, Italy), where a particle accelerator is used for research on hadrontherapy.

---

[61] This is what the authors define 'entitativity'.

3.5 Services provided to third parties

We only briefly mention here that many RIs, particularly in applied research, can provide services to users outside the scientific community. When concrete applications are found for the outputs of fundamental research, a new set of benefits may arise for people who are going to use those outputs. Applied research can contribute to address many societal challenges, for example by finding new ways for ensuring energy security and efficiency, tackling climate change, reducing environmental pollution, mitigating the risk of natural disasters and improving health conditions.

For these benefits it is in principle possible to identify a demand, which relates to the ultimate beneficiaries of the services provided by the RI. For example, for RIs whose objective is to test new methods to address environmental risks (soil erosion, floods, etc.), beneficiaries of applications of new knowledge created at the RI are all people who would enjoy an improvement of social welfare because of the risk reduction. The main method to value these benefits is based on the willigness to pay for a reduced exposure to risks, thus nothing really new is needed in the CBA methods to be used, except perhaps a more sophisticated ex-ante analysis of demand because the services provided may be rather peculiar. In some cases, the avoided cost approach can be applied to value economic benefits: by making available a new technology or product to third parties, the RI allows those actors to save at least part of the development and testing cost. Another example is offered by software for computing or big data management, developed for the purpose of experiments carried out at the RI but which are then released to the public and find application in other fields (e.g. in the finance sector).

Other typical examples of RI services can be found in the health field. Research targeted to improving human health by testing new drugs or new forms of treatment can benefit all those people who will enjoy higher life expectancy or quality thanks to the application of the more advanced and effective therapy.

For instance, at CNAO (National Centre for Oncological Treatment in Pavia, Italy) tens of thousands of patients with some types of solid tumors will be trated over the next decades. This fact is going to have an expected impact in terms of life-years saved suitably adjusted by the quality of life, for which valuation there is a well established CBA approach. Following the literature, the monetisation of an increase in the life expectancy encompasses the estimation of the Value of Statistical Life (VOSL) and the related Value of a Life Year (VOLY). The former is defined as the value that society deems economically efficient to spend on avoiding the death of an undefined individual. The latter represents, instead, a constant value to be attributed to each life year lost due to premature death. Different methods of measuring or approximating society's willingness to pay for reducing the risk of death exist, ranging from contingent valuation survey to benefit transfer, from cost of illness to human capital approach (see, for instance, Viscusi, W. and Aldy, J. E., 2003; Ashenfelter, O., 2006; Sund, B., 2010; OECD, 2012). Once determined the monetary VOSL, in order to quantify the benefit arising from mortality changes, it should be multiplied by the expected number of lives saved thanks to the treatment provided in the RI.

In case of health RIs, unsurprisingly, preliminary empirical testing shows that the magnitude of this *A*-benefit is by far greater than any other item in the Equation (5). In general, for applied research, the *S*, *H*, *C* effects will be contained relative to the *A*-effect (while the *T*-effect is largely variable across fields).

3.6 The pure value of discovery

In the previous sections we have presented a general framework on the way the observable economic use-benefits (B_u) of research infrastructures may be defined and how they could be treated in a CBA framework. We have left aside the B_u term. Nobody can *directly* forecast the full impact of discoveries *per se* on social welfare. Attempting to value them by observing ex-post how they change human life is also pointless, because just as one can mention hundreds of examples of how the verification or falsification of theories has concretely affected human life, one may also mention hundreds of examples of scientific discoveries that have had, until now, no direct tangible socio-economic impact at all.

The avenue we have proposed above is to focus on knowledge outputs as the main direct effect of the RI, and on technological spillovers, human capital accumulation and cultural effects as indirect effects, plus services provided to third parties if any. In most cases for applied research all this is enough to justify a well-designed RI, or in any case to adequately assess its NPV. However, for infrastructures for basic research we would miss an important piece of the puzzle and grossly underestimate the whole impact of scientific enterprises. We argue that there is a residual impact of the RI on social welfare and this is related to its discovery potential.

A RI can discover exactly what it was designed to discover, e.g. the Higgs boson, an extrasolar planet, the change of temperature of the ocean, a virus, etc. This can be labelled as the ‘known unknown’ (Kulmus, 2012). But the infrastructure can also discover something unexpected (the ‘unknown unknown’) or it can discover nothing. The latter in turn may happen for different reasons: either because the experiment falsifies a previous theory (which is of course a valuable result), or because the experiment was not well designed and did not work. All this amounts also to new knowledge, to some extent, of different intensity and meaning. While the value of papers stemming from such discoveries would reflect the social benefit to scientists of advancing knowledge within their community, the discovery itself could bring a number of further improvements on human life.

Synchrotron radiation, for example, has been used to investigate the dynamic process through which water is stored inside rocks within the Earth and returns to the surface in volcanic eruptions or ocean ridges. Experiments succeeded to show how this mechanisms works, thus contributing to produce a more detailed understanding of the role of water inside the Earth. This new information can be used in order to study the effects of global warming on the levels of the oceans (Science and Technology Facilities Council, 2010); for the time being, however, we have no idea of all the concrete results that this discovery may lead.

In principle, an ex-post analysis may reveal some of the intrinsic uncertainty of the RI, but not all of it, because the nature of knowledge is such that the residual effects of a discovery may appear in a very distant future well beyond the decommissioning of the RI. In a sense, any ex-post analysis of a RI is in fact an interim analysis, carried out at a certain moment of the RI’s horizon, but not at its end. Nevertheless, in some cases ex-post evaluation could allow the assignment of a measurable value to the residual effect of the RI associated with the pure or intrinsic value of discovery. Ex-ante, however, the residual is by definition unknown since there is a very large uncertainty both about the probability of materialisation of any discovery and about its possible welfare impact. As we have shown, when we guess that $NPV_{RI} - B_n = NPV_u > 0$, there is no need to go further in the analyis. Strictly speaking one should just assume that B_n is non-negative, and no more is needed. What to do when however $NPV_u < 0$? It is important to stress that NPV_u is only a part of NPV_{RI}. Thus a negative NPV_u does not mean that the society loses with the RI, but it is a signal that we cannot be content with just a guess that B_n is non-negative.

In order to define more precisely this residual element, denoted as B_n in equation (5), we adopt an approach and a terminology borrowed from environmental economics. In the framework of environmental CBA, any good or natural resource can be assigned a total economic value, which in turn can be decomposed into two general classes: use value and non-use value. Use value refers to direct or indirect benefits arising from the *actual use* of an asset (e.g. using a water reservoir for energy production) or its potential or *option use*, indicating the value attached to future opportunities of the goods (e.g. possible recreational use of the water basin). Estimating the option value usually implies that the possible present or future use is already known.

Non-use value denotes the social value for simply preserving a natural resource compared to not preserving it, regardless of its actual or potential (known or unknown) use. Non-use value can be translated into a *bequest value*, arising from the desire to preserve certain resources for the benefit of future generations, or an *existence value* related to knowing that a good (e.g. an animal species in danger of extinction) simply exists even if it has not actual or planned used for anyone and independently of any altruistic motives. There could also be situations in which a practical use of a good can be in principle expected but it is still unknown. In these cases, its value (which can also be referred to the non-use category) is determined by what is generally called '*quasi-option* value'.

This concept was introduced by Arrow and Fisher (1974) when studying how the uncertain effects of some economic activities could be irreversibly detrimental to future environmental preservation. The quasi-option value describes the impact of a development intervention in one period on expected costs and benefits in the next, i.e. the expected net benefits in future periods that are conditional upon the realised benefits in the present period. Elaborating on this, Conrad (1980) highlighted that the notion of quasi-option value is equivalent to the expected value of information. The value of lost and new options allowed by an investment project implemented today is an expected value based on what one might learn. The same interpretation is found in Atkinson *et al.* (2006: 21), who define the quasi-option value as the 'difference between the net benefits of making an optimal decision and one that is not optimal because it ignores the gains that may be made by delaying a decision and learning during the period of delay' and, we may add, the unknown losses that may occur by delaying the same decision.

Retaining this terminology, we can assert that the RI's main benefits – creation of knowledge outputs, technological externalities, human capital accumulation and cultural impact of the outreach, and services – capture the use value of RI projects for different categories of stakeholders, from scientists, to firms, students and the general public, and beneficiaries of other services. The value of unknown non-use vale effects, on the other hand, is determined by a sum of quasi-option and existence values. More specifically, we argue that a RI has a quasi-option value in the sense that it could generate discoveries that would produce positive impacts that cannot be estimated at the time when the funding decision is taken. Also, an existence value can be attributed to RI's discoveries, referring to the intrinsic value of knowing the object of the discovery, regardless of the fact that it might find some use soon or later.

Using these concepts, originally conceived for environmental goods and natural resources, to value also other categories of goods is not a new practice. Arrow and Fisher (1974: 319) conceded that the quasi-option value is a general notion that may be applied outside of environmental economics, as it is linked to uncertainty, information and irreversibility issues affecting decision making in general. Existence value is often discussed when attempting to assess the value of culture, arts or sport, to which an intrinsic but immaterial value is attached. Some people get value from the existence of a cultural good or service, despite not using or engaging with it, for example because of the pride they feel towards a local cultural organisation or the importance attached to the existence of heritage, despite it not being a subject of direct interest to them (DCMS, 2010: 23-24). Similarly, a scientific discovery could benefit people who have preference for knowledge. We are not referring here to scientists, but to 'ordinary' people who, even if do not fully grasp the meaning and implications of a discovery, are happier simply because that discovery occurred.

Quasi-option value and existence value are two distinct concepts. First of all, if quasi-option value could be either positive or negative, producing either an increase or a decrease of social welfare, existence value can always be regarded as intrinsically positive, or at least nil: people can be expected to be better off with the discovery, or completely indifferent to it, but in general an increase in knowledge *per se* does not reduce social welfare. Secondarily, the quasi-option value for the unknown effects of a discovery is completely uncertain, and thus no preferences can be imputable to it, as long as the effects remain unknown. Instead, people can have some preferences about a good's existence; they are unlikely to have preferences if they do not know or understand the issue at stake, but if they get "some" information it is reasonable to assume that preferences will arise, allowing them to choose between two states of the world: i.e. one in which the scientific discovery occurs and one in which it does not. Indifference is also possible, and it is unrelated to the impossibility to choose that is instead related to quasi-option values. Thirdly, and very much related to the previous aspect, the lack of any preference for quasi-option value entails the impossibility to derive a demand for those unknown goods and, hence, to measure such a value. On the contrary, preferences for the existence value of goods (discoveries) in principle give the possibility to guess ex-ante this further component of the total economic value of research.

This means that we can decompose the residual value B_n of the RI CBA model, into two separate components: the quasi-option value $QOV$and existence value EXV of the discovery. Both quasi-option and existence value are *instantaneous* variables expressed at time $t = 0$, therefore they do not need to be discounted and predicted over time. We have that:

$$B_n = QOV_t + EXV_t \tag{10}$$

where $QOV_t \in (-\infty, +\infty)$ and $EXV_t \in [0, +\infty)$

and $t = 0$, i.e. the time when the evaluation is performed. As from our discussion above, we suggest that while the quasi-option value term remains completely unknown ex-ante (and ex-post for a long time), some empirical analysis about the existence value's size could be made.

The standard way of estimating non-use values for which no observable price system exists is to use stated preferences techniques, i.e. techniques based on answers given by a representative sample of the population of interest to derive respondent's tacit willingness to pay for a good.[62] On the same vein, one could attempt to grasp the willingness to pay of taxpayers for having the RI compared to not having it, regardless its actual or potential use. As mentioned elsewhere, for scientists who understand the issue at stake, the willingness to pay is probably high and is already captured by writing, reading, citing and presenting papers. But what can we say about a median tax-payer, who has limited information about the topic? Conceptually the issue is not different from estimating the willingness to pay for climate change policy or for conservation of bio-diversity. Most people have only very vague ideas about these issues, but, when – to a certain extent – they are given information, they express their attitudes in appropriately designed surveys.[63]

Attempts to measure the existence value of goods in practice have been made through contingent valuation or other techniques. Contingent valuations include choice experiment or conjoint analysis methods, previously mentioned when addressing the issue of the value of RI's cultural effects (Section 3.4). Contingent valuation has been developed as a method for eliciting market valuation of damages to environmental resources, but it has also been used to value a wide range of non-market goods and services, such as museums (Tohmo, 2004), cultural heritage (Willis, 1994; Tuan and Navrud, 2008), local football clubs (Barlow, 2008), and others.[64] Other methods include revealed preference techniques, which assume that the existence value can be determined through the observation of economic behaviours in a related market, such as voluntary contribution to organizations devoted to the preservation of a public good (animal species, wilderness areas, etc.); or the so called 'Wellbeing Valuation' approach, based on estimating monetary values for non-use values by looking at the way a

---

[62] We have already mentioned the use of contingent valuation to derive the willingness to pay for virtual cultural activities and services provided by the RI (Section 3.4).

[63] Some criticism has been raised against the legitimacy of non-use values (Weikard, 2005; Boudreaux et al., 1999), according to which only goods with a use-value have an economic value and impact on individuals' utility function. However, the importance of existence value as a component of the total economic value of goods is in fact unanimously advocated by contingent valuation studies and even reflected in some legislation, e.g. in the United States Federal Preservation Regulation (Dana, 2004).

[64] For example, Jura Consultants (2005) estimated that the museum, library and archive services of the community in Bolton (UK) was worth £10.4 million, of which £3 million related to non-use value.

good impacts on a person's well-being, and finding the monetary equivalent of this impact[65] (for an example of application in the culture sector, see Fujiwara, 2013).

Consistently with environmental and culture economics, we maintain that an existence value can be attributed to the potential discoveries of research, and that some efforts to incorporate existence values in CBA can be attempted, keeping into account the methodological issues discussed above. In particular, the way how we propose to value the discovery's existence value is in line with Boardman *et al.* (2006: 229):

> 'Should existence values be used in CBA? The answer requires a balancing of conceptual and practical concerns. One the one hand, recognizing existence values as pure public goods argues for their inclusion. On the other hand, given the current state of practice, estimates of existence values are very uncertain. This trade-off suggests the following heuristic: *Although existence values for unique and long-lived assets should be estimated whenever possible, costs and benefits should be presented with and without their inclusion to make clear how they affect net benefits*. When existence values for such assets cannot be measured, analysis should supplement CBA with discussion of their possible significance for the sign of net benefits.'

Recalling the CBA model introduced in Section 2 we have now the following:

$$NPV_{RI} = NPV_u + B_n = (PV_{B_u} - PV_{C_u}) + (QOV_0 + EXV_0). \qquad (11)$$

The measurable net present value could be either higher, lower than or equal to zero, when the residual effects are not measured. However, the residual effect at time zero (observation time) consists of one component which is completely unknown and which could assume both positive and negative values (the quasi-option value of discovery), and another component which could take (non-strictly) positive values and is related to the existence value of a possible discovery, whose measurement is however not obvious.

We claim that in practice, the negative range of possible unknown values for QOV should simply be ignored, i.e. assumed to be nil, except in extreme cases of potentially very dangerous research, and hence we could be satisfied if use-benefits and the existence value of possible discovery more than counterbalance costs. This gives us the possibility to look at the CBA test from another perspective and to state whether a project is socially desirable by testing the following hypothesis when $NPV_u < 0$ (in the other cases, as mentioned, this test is not necessary):

$$B_n \cong EXV_0 > -NPV_u. \qquad (12)$$

This means that the net present value of the residual effects of a RI, proxied by the existence value of discovery, should be greater than the complement of measurable negative net present use-value. In other terms, the RI is deemed socially beneficial if the (positive) existence value is greater than the net (i.e. negative NPV of measurable use-components) costs.

Having set a conceptual frame, let us turn to possible strategies for the empirical estimation of EXV. First of all, it is helpful to think to this problem in the perspective of the median tax-payer. If we assume, for example, that the net present value of a government-owned RI without considering the EXV related to potential discovery is (–*x*) million discounted Euro, this would be the social cost (at shadow prices) of building the RI. Dividing this value by the number of taxpayers, or, in general, by the total population with a stake in the project, gives the per capita *minimum per capita willingness to pay for the RI's existence value*.

---

[65] In practice, a survey is submitted to measure the effect of a public good on happiness. If a good increases happiness by 1 index point per year and additional x euro of income also increases happiness by 1 index point, then the equivalent value of the public good is x euro.

This heuristic approach for project valuation entails two separate steps: first the quantification and valuation of use-benefits and costs, and, second, the estimate of the minimum amount of money that taxpayers should be asked to pay for the existence value of a discovery.[66] Having set this minimum threshold, we can proceed in three ways: stated preferences techniques, revealed preferences techniques and benefit transfer.

Under the first approach, a contingent valuation on a representative sample of tax-payers should test the willingness to pay an amount of money equal or greater than the threshold necessary to get a positive NPV_{RI}. We can see this as a formalisation of the way scientists and policymakers often implicitly justify public spending based on guesses of social preferences for non-standards goods. The questions & answer online site of the Earth Observation Environmental Satellite (Envisat),[67] replies to a question on the cost of the infrastructure, by stating:

> 'Envisat cost 2.3 B Euro (including 300 M Euro for 5 years operations) to develop and launch (launch price tag: 140 M Euro). This is equivalent to 7 Euro per head of population across all the ESA member states, or about one cup of coffee per year spread over its 15 year lifecycle.'

One possible objection to our suggestion is that the sampling and the information provided in the contingent valuation exercise should make it rather costly. While this is an important consideration, we do not see why obfuscating the true cost of large scale scientific enterprises to the tax-payers should be advocated as a better approach. Moreover, we are not convinced that the typical cost per capita of a well-designed contingent valuation would more than a very modest fraction of the overall cost of the RI in the first place, particularly for the large ones. Another possible objection is that asking individuals their willingness to pay for the mere existence of any good may not be easy and may result to be biased by a number of individual, cultural and socio-economic circumstances (Carson and Groves, 2007; Carson, 2012). In order to address these issues, the evaluator can take into account a number of recommendations developed since the early Nineties by a panel of distinguished economists[68] for the US National Oceanographic and Atmosphere Agency (NOOA, 1993), including indications about the modalities and structure of the interviews.[69]

As a second approach, in order to overcome the difficulty of explicitly stating a willingness to pay in a situation where it is not possible to observe markets for science and thus to take from them some hints on the pure value of discovery, valuation methods based on revealed preference can be conveniently employed. The social value attributed to the existence of research projects by population can be revealed by data on donations. Health research, for example, is widely supported by voluntary donations. In some countries taxpayers can name a charity to whom a percentage of their taxable income is donated and several scientific institutions are supported in this way.[70] Universities regularly receive donations for research by firms and by individuals. All this shows that a generic benchmark about the willingness to pay for science (and even for different fields) can be revealed by observation of actual behaviour of very large numbers of individuals.

A third approach, not necessarily alternative to the previous ones, would be to recur to benefit transfer. In this case instead of sampling respondents, a meta-analysis of contingent valuation studies on the existence value of goods produced by other projects is used to establish a benchmark median value or a range of values. Then the minimum per capita value that the EXV of the RI should take to compensate for the negative net use-costs can be compared with such benchmark. If it is well within

---

[66] At least for expected discoveries of the 'known unknown' type (see above). More difficult would be to state a willingness to pay for the 'unknown unknown'.

[67] http://www.esa.int/Our_Activities/Observing_the_Earth/Envisat_FAQs/(print).

[68] K. J. Arrow and R. M. Solow among them.

[69] For some reflections about the applicability of the NOAA recommendations to the field of scientific research infrastructures, see Florio 2014b.

[70] In Italy, this amounts to 0.5% of personal annual income.

the range, or in the median to lower bound of it, we can guess that the project is as beneficial as other goods for which empirical analysis of an existence value is available.

We conclude that, once the concept of the pure value of discovery is introduced, and even if the quasi-option value in terms of utility remains unknown, the NPV of RI projects in basic research can be evaluated. There is a certain advantage if the decision process is explicitly based on the evaluation of the project net present value, distinguishing between use and non-use benefits and costs. In this way, the focus of the decision could be shifted to the assessment of use-benefits and costs versus the ‘residual’ unknown benefit related to the existence value of the discovery. Results can then be discussed and used in the decision making process to compare costs of similar but alternative projects producing similar outputs. This is of obvious practical importance. If the ‘one cup of coffee per year’ test is passed for many projects, but not for the one under study, clearly the qualitative case for saying that there is a net benefit for the society would need to be much stronger than otherwise. Possibly a specific well-designed contingent valuation survey would be mandatory in such circumstances if the project needs to be supported by the tax-payer. We discuss some related topics for empirical analysis in Florio (2014b).

4 Risk assessment

Once all costs and benefits entering the CBA model of research infrastructures have been identified and their baseline values have been estimated, an assessment of the probability of error related to each estimate might still appear a daunting task, particularly from an ex-ante perspective.

As the empirical evidence suggests, similar problems arise when forecasting the project investment costs. Flyvbjerg *et al.* (2003) stress that, when financing large scale public infrastructure projects, there exists a tendency to underestimate costs, overestimate revenues, undervalue environmental impact and overvalue economic development effects. This occurred for example in the case of the Superconducting Super Collider project, with an estimated investment costs (at nominal prices) which passed from USD 4.4 billion in 1986 to 11 billion in 1993, or the National Ignition Facility, whose cost was expected to be around USD 2.07 billion in 1995 and almost twice as much (USD 3.89 billion) five years later (including the cost of construction and experimental programmes).[71]

Nevertheless, no RI project would be implemented if there was not a widespread belief that some expected scientific results are likely to occur in future and the cost would not exceed a given threshold. There may be extreme cases where these beliefs are shared by a very small set of people, such as a committee of experts, or a prime minister and his advisors, but usually a larger community of scientists and other stakeholders are involved in the decision process. While scientists may be totally unaware of any formalized risk assessment procedure, RIs are designed by highly qualified scientists and engineers who have expectations based on their previous knowledge. In this sense, the riskiness of different variables of the project can be, to some extent, guessed.[72]

For example, in the Shuttle missions, there are probabilities involved in the actual working of different components. Thus, on the technical side, there are risks that are generally taken into account and computed. In this regard, it is interesting to note that the ‘black holes issue’ and the ‘strangelet risk’ at the LHC and Brookhaven have also been dealt by special panels of scientists in terms of probabilities of events and associated catastrophic risks (Maiani, 2012). These examples show that when this is considered necessary, scientists can express their beliefs in terms of probabilities of events. There is

---

[71] The National Ignition Facility, built in California (USA) and operative since 2009, is used to simulate the thermonuclear conditions created in nuclear explosion, serving the research on the behaviour of nuclear weapons without using explosive testing (GAO, 2000).

[72] Knight (1921) refers to risk as ‘measurable uncertainty’, since it can be embedded into a stochastic model by means or probability distributions.

for sure less knowledge on the probability distribution of benefits, also because there is a limited experience with their identification and valuation, but guesses and bets are unavoidably part of the scientific game.

Following the CBA tradition and practice, risk assessment has to be taken into account when designing a project and forecasting costs and benefits. Risk assessment involves the set of qualitative and quantitative methods and procedures aimed at evaluating the probability that a given project will achieve a satisfactory performance.[73] The set of procedures for overall risk assessment is traditionally split into three steps. First, a sensitivity analysis has to be carried out: the impact of each variable entering the analysis of the outcome (the net present value) is assessed by changing each 'best guess' value in absolute terms or by arbitrary percentages, one by one; having then set a criterion to decide whether the variation in the output is sufficiently large, the most critical variables for the CBA can be identified.

To each of them, as a second step, a specific probability distribution function is assigned. Probability distributions are highly dependent on the specific type of project under evaluation and they may be determined from various sources of information, including experimental data, distributions found in literature and adopted in projects similar to the one under assessment, time-series or other sorts of historical data (Vose, 2008).

Third, the project's riskiness is assessed through Monte Carlo simulation techniques, which allow for an estimate of the integral corresponding to the probability distribution function of the net present value, by drawing (without replacement) one value of each critical variable from the respective cumulative distribution function; the extracted values are plugged into the CBA model and the associated NPV is computed. This process if repeated over a large number of iterations. Overall, the usefulness of the Monte Carlo approach is ultimately linked to the fact that, through the law of large numberswhich implies the convergence of the NPV empirical distribution to its 'true' counterparts,the CBA result can be considered in probabilistic terms and the minimum, maximum, mean values and standard deviation of the NPV can be computed.

We suggest to deal with the different types of costs and benefits of the RI project separately and to check the possibility to get expert opinions about the probability of occurrence for each of them. Carrying out the risk analysis would imply an estimate of the probability distribution of at least the following, certainly critical, variables: quantities and shadow prices of capital and operating expenditure on the cost side and knowledge, technology, human capital formation and outreach on the benefit side.

More specifically, the distribution of knowledge-related variables can be reduced to a range of number of papers (p_{it}, to use the notation introduced in Annex 2), citations (q_{it}) and references (k_{it}) of scientific literature, possibly based on previous experience in scientometric analysis, and to a range of values for the marginal social cost of each of such outputs (considering the shadow cost of paper production ζ_{it} and of citations ξ_{it}). As for technological spillovers and human capital accumulation, the incremental profit experienced by companies of the RI's supply chain (Π_{jt}) or the incremental annual earnings of former RI's students or workers (I_{zt}) are certainly critical variables for which a probability distribution should be assumed. The same is true for estimates of the willingness to pay of the general public for cultural effects associated to the RI (W_{gt}). Assessing the risk attached to the net present value of the RI project also enables a determination of the range of variation of the notional existence value of discovery (*EXV*). Actually, as we argued in Section 3.6, the latter can also be defined as the minimum value of discovery that would switch a negative NPV into a positive one (this in fact would be a simple sensitivity analysis, but contingent valuation and other empirical approach would give a range of values).

---

[73] See more details on the procedures for the risk assessment in Florio (2014a).

Nothing can be said, however, about 'serendipity' effects (the World Wide Web being a notable example) and unpredictable accidents (the helium leak occurred in the LHC magnets in 2008 is perhaps as another example), as well as more in general the quasi-option value of possible discoveries, which are unpredictable by definition and therefore remain uncertain.

In addition to the riskiness attached to variables entering the social CBA either on the cost or on the benefit side, the results of the CBA model for RIs could be strongly influenced by two important parameters that we cannot discuss here: the time horizon[74] of the analysis and the social discount rate.[75] Our recommendation is that one should assess the variation of the CBA results also subject to different assumptions on the length of the time horizon ($\mathcal{T}$), which would most likely affect the size of the benefit of knowledge creation, and of the chosen discount rate (r) and discounting function (exponential, hyperbolic or others).

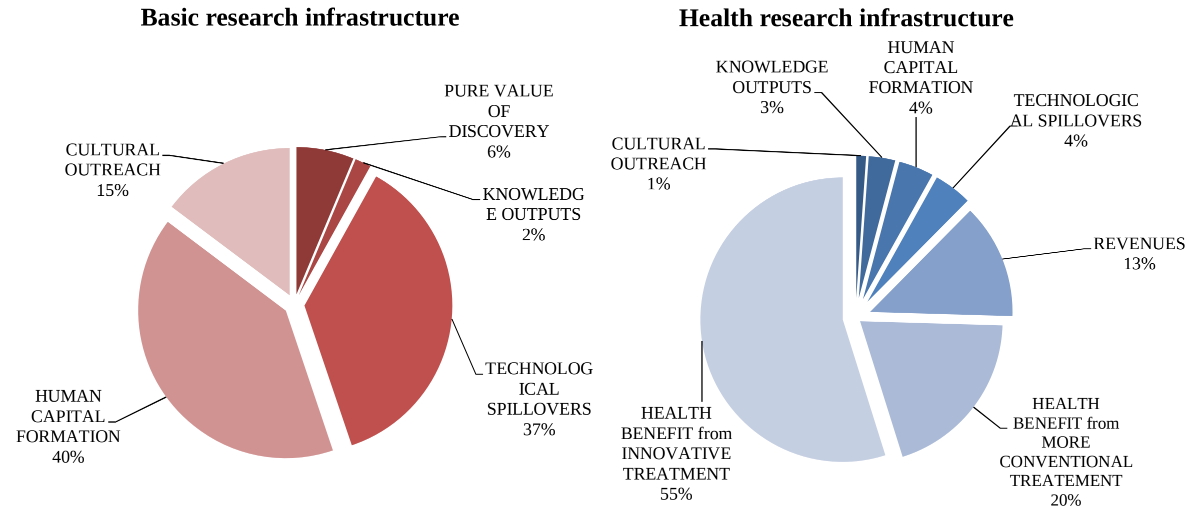
In a purely illustrative way, we show in Figure 1 a possible distribution of the discounted benefits arising from two infrastructure, one for basic research, the other one for applied research, specifically in the health sector. Figure 2, on the other hand, shows some possible probability distribution functions associated to potentially critical variables plugged into the CBA model such as the investment costs, the incremental salary and the profit margin for suppliers as well as the resulting probability distribution function of the NPV.

---

[74] What is peculiar of research infrastructures as compared with projects in other sectors is not the long-term construction or operation timeline, but their very long and even permanent benefits. The accumulation of knowledge, which is the core RI's direct benefit has a longer, possibly infinite, time horizon than the period of operation of the RI. However, it should be realized that an infinite time horizon would lead to a paradoxical result, because any large investment cost spread in a finite range of years would be less than the sum of any small benefit spread in an infinite time horizon, whatever non-strictly positive social discount rate. Therefore it seems reasonable to assume a long, but finite, time horizon for the benefits of a RI. Another reason why the life cycle of a RI should be taken as finite is related to the obsolescence process of the value of knowledge over time, which is observable, for example, in the temporal trend of citations to the results of experiments carried out at the RI. In addition to knowledge creation, other effects of RIs could continue after the infrastructure's decommissioning: namely, the human capital accumulation and technological spillovers. As discussed in Section 3.3, the former benefit, captured by the incremental salary of students who have been trained at the RI, could be as long as their work career, at most. As far as technological externalities are concerned, measured as incremental profits of the RI's supplier firms which enjoyed learning-by-doing effects and the whole expected profit of spin-offs, in principle they could extend for the whole life of the company. Hence, unlike knowledge creation, human capital and technological spillover effects usually have an end point. Even if this is beyond the CBA's horizon, the residual value of these effects can be included in the final year of the analysis.

[75] The social discount rate (SDR) expresses the rate at which society is willing to postpone a unit of current consumption in exchange of more future consumption. In most of CBA practice, a constant discount rate is used, which implies an exponential discounting process of the project's inflows and outflows. However, evaluating projects with impacts accruing for many years or decades in the future poses a dilemma on the choice of the SDR. With a constant rate, benefits occurring far in the future are discounted more than costs of investments, which instead usually take place in the initial years of the time horizon. This could lead to a negative benefit/cost ratio for the project and to the decision of not implementing it, disregarding however that the same project could have tremendous benefits on the welfare of future generations. A possible way to deal with this problem is to adopt a sufficiently low discount rate, as done for example in the Stern Review on the Economics of Climate Change (HM Treasury, 2006). Another possibility suggested in the literature is to use a SDR that declines over time, to take into account that in facing the decision between a smaller reward soon and larger reward later, individuals would apply a lower discount rate in the long run.

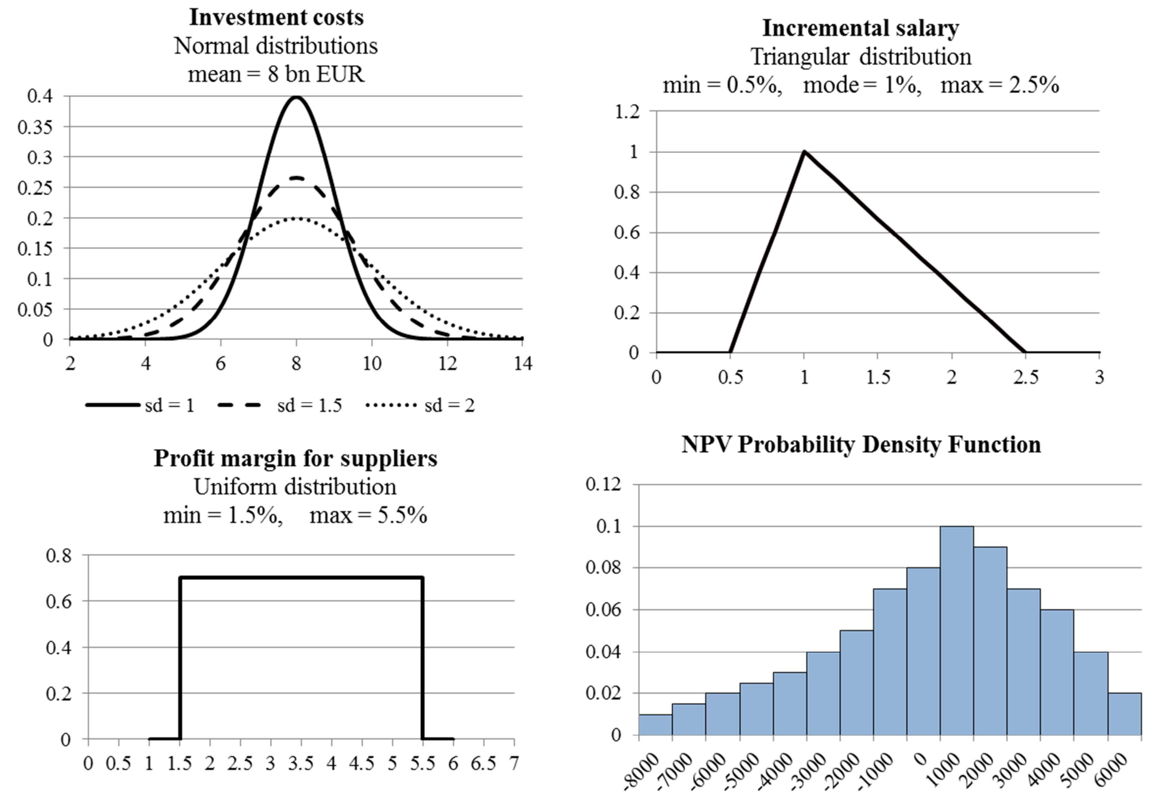
Figure 1 Example of distribution of the discounted benefits for two hypothetical research infrastructures



Note: The share of benefits related to knowledge outputs does not include the social value of publications produced by the RI internal scientists, which cancels out with the scientific personnel cost. Thus it includes only the benefit associated to citations and papers produced by outsider scientists.

Source: Authors

Figure 2 Examples of probability distribution functions of critical variables and the NPV



Source: Authors

5. Conclusion: the socio-economic NPV of research infrastructures

Our approach to the social CBA of large-scale, capital intensive, scientific projects can be summarized as follows. A research infrastructure is, after all, an infrastructure, and we need to exploit as far as possible our understanding of the generic economics of infrastructure projects, on which there is a wide literature (see Picot *et al*., forthcoming). This is particularly important on the cost side, where issues such as project delimitation and cost apportionment must be solved.

On the benefit side, research is a service provided to society. As for any other service, it is crucial to estimate its demand, in a context where there is no market mechanism shaping and rationing demand. We have suggested that the direct notional demand for RIs comes from scientists and that a peculiarity of research infrastructures is that their users are also service producers. This implies an estimation of social benefits of knowledge outputs based either on the marginal value that scientists would be willing to pay for working on the RI project, regardless the actual monetary compensation they get, or on the marginal cost of knowledge production, valued at a suitable shadow price. It is unknown whether the two things are actually equal, or if the shadow price of knowledge is a combination of the two notions. Interestingly, however, in both cases, costs and benefits cancel out, so that, in a sense, the RI in part pays for itself. If the scientific personnel marginal cost for the RI is also taken as a marginal benefit, as we suggest in Section 3.1, in order to pass the CBA test the present value of investment costs and of other technical costs must be less than the present value of indirect knowledge outputs, other positive benefits and eventually uncertain 'discovery' residual effects of the project.

The investment cost and operating costs (excluding scientific personnel) are relatively easy to be computed. The core benefit assessment is then based on six main dimensions: the social value of indirect knowledge outputs, technological spillovers based on free access to new knowledge and learning-by-doing, the increase of human capital, wider cultural effects through outreach activities, services provided, and the pure value of discovery.

The magnitude of each of the first group of five use-effects significantly varies depending on the size of the RI, its field of research, the kind of activity mainly carried out (either fundamental or applied research) and external factors. Among these, the absorption capacity of suppliers to leverage the learning acquired by working for the RI project is crucial. A proactive, long-term approach to leveraging the spillover potential would be needed in order to maximise the benefits of collaboration between the RI and the industrial sector. A similar challenge arises with the benefit from human capital accumulation, which is bigger if the national institutions and firms that employ former RI-students succeed at making good use of individual learning and capacity, but also if the skills acquired during the training period at the RI are not so specific that could hardly be transferred in another context.

Additionally, applied research infrastructure might produce other types of benefits on the final users of the service, which are not already incorporated in the aforementioned benefits, when research is linked to services to beneficiaries outside the scientific community (e.g. patients of health research facilities, or the general public for meteorological or environmental scientific projects including advanced monitoring techniques).

All these variables should be expressed in terms of expected values, to be estimated through a probabilistic risk analysis: this implies assuming certain distribution of quantities and shadow prices of capital costs, operating costs and the RI's core benefits.

This analysis of use-effects may miss however a substantial part of the story: the potential unknown effects of scientific discovery. The net present value of this residual effect should be duly compared at least qualitatively to the estimate of NPV of use-effects. Its valuation could also be attempted by means of a contingent valuation aimed at assessing the willingness to pay of stakeholders who would fund, either directly or indirectly (through taxes), the RI project. To what extent such a residual effect could actually be captured by the CBA of the investment and to what extent the CBA can reveal to be a useful decision tool in this area is an open question, unexplored by earlier literature, that can be answered only through a serious attempt to study actual data of some RI projects.

The CBA model for research infrastructures that has been discussed throughout this paper is offered as a starting point for empirical testing and further conceptual analysis summarized. Annex 3 summarizes it in a compact form.

References

Abt, H.A. and Garfield, E. (2002). ‘Is the Relationship between Numbers of References and Paper Lenghts the Same for All Sciences?’, *Journal of American Society for Information Science and Technology*, 53(13): 1106-1112.

Acs, Z.J., Audretsch, D.B., Feldman, M.P. (1992). ‘Real effects of academic research: comment’, *The American Economic Review*,92 (1), 363–367.

Adams, J. (1990). ‘Fundamental stocks of knowledge and productivity growth’, *Journal of Political Economy*, Vol. 98(4), pp. 673-702.

Arrow, K. (1962). ‘The Economic Implications of Learning by Doing’, *The Review of Economic Studies*, Vol. 29 (3): 155-173.

Arrow, K. J., and Fisher, A. C. (1974). ‘Environmental Preservation, Uncertainty, and Irreversibility’, *The Quarterly Journal of Economics*, 88 (2): 312-19.

Ashenfelter, O. (2006). ‘Measuring the Value of a Statistical Life: Problems and Prospects’, Discussion Paper Series No. 1911 , Institute for the Study of Labor, Bonn, Germany.

ASVAPP – Associazione per lo Sviluppo della Valutazione e l’Analisi delle Politiche Pubbliche, (2012). *Counterfactual impact evaluation of Cohesion Policy: impact and cost-effectiveness of investment subsidies in Italy*, Final Report to DG Regional Policy prepared for the study ‘Counterfactual Impact Evaluation of Cohesion Policy. Work Package 1: Examples from Enterprise Support’.

Atkinson, G., Mourato S., Pearce D.W. (2006). *Cost-Benefit Analysis and the Environment. Recent developments*, Paris: OECD Publishing.

Autio E., Bianchi-Streit, M. and Hameri, A-P. (2003). *Technology transfer and technological learning through CERN’s procurement activity*, CENR 2003-005, CERN Geneva.

Autio, E., Streit-Bianchi M., Hameri A-P. and Boisot, M. (2011). ‘Learning and Innovation in procurement: The Case of ATLAS-Type Projects’, in M. Boisot, M. Nordberg, S. Yami and B. Nicquevert (eds.) *Collissions and Collaboration. The Organization of Learning in the ATLAS experiment at the LHC*, Oxford: Oxford University Press.

Bacchiocchi, E. and Montobbio, F. (2009). Knowledge diffusion from university and public research. A comparison between US, Japan and Europe using patent citations. *The Journal of Technology Transfer*, Vol. 34(2): 169-181.

Bach, L. (2013). ‘Developing a new approach for evaluating the impacts of RIs: Synthesis of the main results of EvaRIO’, presentation at the conference Evaluation of economic impacts of research infrastructures, Strasbourg, October 16-17.

Baggott, J. (2012). *Higgs – The Invention and Discovery of the ‘God Particle’*, Oxford University Press.

Barlow, A. (2008). ‘Do we know the true value of football? A review of methodologies used to value public goods’, Working paper 3, Substance.

Baum, W.C. and Tolbert, S.M. (1985). *Investing in Development: Lessons of World Bank Experience*, Oxford University Press (USA).

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, CERN, Genève.

Blaug, M. (1987). 'Book Review of Economics of Education: Research and Studies', *Journal of Human Resources*, 24 (2): 331-35.

Boardman, A.E., Greenberg, D.H., Vining, A.R. and Weimer, D.L. (2006). *Cost-Benefit Analysis: Concepts and Practice*, 3rd edn, Upper Saddle River, New Jersey: Pearson Prentice Hall.

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

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

Boisot, M., Hoffmann H.F. and Nordberg M. (2011). 'ATLAS and e-Science', in M. Boisot, M. Nordberg, S. Yami and B. Nicquevert (eds.) *Collisions and Collaboration. The Organization of Learning in the ATLAS experiment at the LHC*, Oxford: Oxford University Press.

Boisot, M., Jenni, P. and Nordberg, M. (2011). 'What is ATLAS?', in M. Boisot, M. Nordberg, S. Yami and B. Nicquevert (eds.) *Collisions and Collaboration. The Organization of Learning in the ATLAS experiment at the LHC*, Oxford: oxford University Press.

Boudreaux, D.J., Meiners, R.E. and Zywicki T.J. (1999). 'Talk is Cheap: The Existence Value Fallacy', *Environmental Law*, 29(4): 765-802.

Broad, W.J. (1990). 'Big Science: is it worth the price? A periodic look at the largest new research projects. Heavy Costs of Major Projects Pose a Threat to Basic Science', *The New York Times*, published on May 27, 1990.

Camporesi, T. (2001). 'High-energy physics as a career springboard', *European Journal of Physics*, Vol. 22: 139-148.

Caplin, A. and Leahy, J. (2004). 'The Social Discount Rate', *Journal of Political Economy*, 112(6): 1257-1268.

Card, D. (1999). 'The causal effect of education on earnings', Chapter 30 in O. Ashenfelter and D. Card (eds), *Handbook of Labour Economics*, Volume 3, Elsevier Science.

Carrazza, S., Ferrara A. and Salini S. (2014). 'Research infrastructures in the LHC era: a scientometric approach draft paper produced in the frame of the research project "Cost/Benefit Analysis in the Research, Development and Innovation Sector", part of the EIB University Research Sponsorship programme (EIBURS).

Carson R.T. (2012). 'Contingent valuation: a practical alternative when prices aren't available', in *Journal of economic perspectives*, vol. 26, 4, Fall, 27-42.

Carson R.T. and Groves T. (2007). 'Incentive and informational properties of preference questions', in *Environmental and Resource Economics*, Vol.37, Issue 1, pp 181-210.

CERN (2009). 'LHC the Guide – FAQ', *CERN-Brochure-2009-003-Eng*, CERN Communication Group.

CERN (2011). *CERN Communication strategy 2012-2016*, CERN/DG/CO/2011-1, 12 October 2011, Draft 5.

CERN (2013). *CERN Annual Report 2012*, prepared by the CERN Communication Group, Geneva.

CERN (2013). *Knowledge Transfer 2012, Brochure 2013-001*, Knowledge Transfer Group, Communication Group, Genève.

Chyi, H.I. (2005). 'Willingness to Pay for Online News: An Empirical Study on the Viability of the Subscription Model', *Journal of Media Economics*, 18(2): 131-142.

Cohen, W.M. and Levinthal, D.A. (1990). 'Absorptive Capacity: A New Perspective on learning and Innovation', *Administrative Science Quarterly*, Vol. 35 (1): 128-152.

Conrad, J.M. (1980). 'Quasi-Option Value and the Expected Value of Information', *The Quarterly Journal of Economics*, Vol. 94 (4): 813-820.

COST Office (2010). *Benefits of Research Infrastructures beyond Science. The Example of the Square Kilometre Array (SKA)*, Final Report.

Cowan, R. and Zinovyeva, N. (2013). 'University effect on regional innovation', *Research Policy* 42: 788-80.

Crosbie, V. (2002). 'Online content: The 2002 report', available at http://www.clickz.com/clickz/column/1699449/online-content-the-2002-report, Clickz.com.

Czech Ministry of Education, Youth and Sport and JASPERS (2009). *Background methodology for preparing feasibility and cost-benefit Analysis of R&D infrastructure projects in Czech Republic*, supported by the Cohesion Fund and the European Regional Development Fund in 2007-2013.

Dana, D.A. (2004). 'Existence value and Federal Preservation Regulation', *Harvard Environmental Law Review*, 28: 343-399.

DCMS – Department for Culture, Media and Sport (2010). *Measuring the value of culture: a report to the Department for Culture Media and Sport*, prepared by O'Brien D.

De Solla Price, D. J. (1970). 'Citation Measures of Hard Science, Soft Science, Technology, and Non Science', in C.E. Nelson and D.K. Pollack (eds) *Communication among Scientists and Engineers*. Lexington, MA: Heath, 3-22.

De Solla Price, D.J. (1963). *Little science, big science*. New York: Columbia University Press.

Del Bo, C. (2014). 'The rate of return of investment in R&D infrastructure: an overview', draft paper produced in the frame of the research project "Cost/Benefit Analysis in the Research, Development and Innovation Sector", part of the EIB University Research Sponsorship programme (EIBURS).

Drèze J.P. and Stern N. (1987). 'The theory of cost-benefit analysis', in Auerbach A.J. and Feldstein M. (eds), *Handbook of public economics*, North Holland, Amsterdam, Vol. II.

Drèze, J. and Stern N. (1990). 'Policy reform, shadow prices and market prices', *Journal of Public Economics*, 42 (1): 1-45.

ESFRI – European Strategy Forum on Research Infrastructures (2011). *Strategy Report on Research Infrastructures - Roadmap 2010*.

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

European Investment Bank (2013). *The Economic Appraisal of Investment projects at the EIB*, Luxembourg.

Feldman, M.P. and Florida, R. (1994). 'The geographic sources of innovation: technological infrastructure and product innovation in the United States', *Annals of the Association of American Geographers*, 84, 210–229.

Florio, M. (2014a). *Applied Welfare Economics: Cost-Benefit Analysis of Projects and Policies*, London: Routledge.

Florio, M. (2014b). *On the existence value of scientific discovery: a note*. Mimeo, Milan European Economy Workshop 12-13 June, 2014.

Flyvbjerg, B., Bruzelius, N. and Rothengatter, W. (2003). *Megaprojects and Risk: An Anatomy of Ambition*, Cambridge (U. K): Cambridge University Press.

Fujiwara, D. (2013). *Museums and happiness: The value of participating in museums and the arts*, The happy museum project, available at http://www.happymuseumproject.org/wp-content/uploads/2013/04/Museums_and_happiness_DFujiwara_April2013.pdf.

Gadd, H., Hansson, G. and Månsson J., (2009). 'Evaluating the impact of firm subsidy using a multilevel propensity score approach', Working paper series n. 3, Centre for Labour Market Policy Research (CAFO), School of Management and Economics, Vähjö University.

GAO – United States General Accounting Office (2000). 'National Ignition Facility. Management and Oversight Failures Caused Major Cost Overruns and Scheduled Delays', Report of the Committee on Science, House of Representatives. Available at: http://www.gao.gov/archive/2000/rc00141.pdf.

Giudice, G.F (2010). A *Zeptospace Odissey, A Journey into the Physics of the LHC*, Oxford University Press.

Gramlich, E.M. (1994). ‘Infrastructure Investment: A Review Essay’, *Journal of Economic Literature*, Vol. 32 (3): 1176-1196.

Griliches Z. (1980). ‘Returns to R&D expenditures in the private sector’, in J. W. Kendrick and B. Vaccara (eds), *New Developments in Productivity Measurement and Analysis*, Chicago: Chicago University Press.

Hagström, W. (1965). *The Scientific Community*, New York: Basic Books.

Hall, B. H., Branstetter, L., Crépon,B., and Mairesse,J. (1999). ‘Does Cash Flow Cause Investment and R&D: An Exploration using Panel Data for French, Japanese, and United States Firms in the Scientific Sector’, in Audretsch, D., and A. R. Thurik (eds.), *Innovation, Industry Evolution and Employment*, Cambridge, England: Cambridge University Press.

Hall, B. H., Mairesse J. and Mohnen P. (2009). ‘Measuring the returns to R&D’, in B.H. Hall and N. Rosenberg (eds.) *Handbook of the Economics of Innovation*, Elsevier, Amsterdam.

Hall, B.H., Griliches, Z. and Hausman, J.A. (1986). ‘Patents and R&D: Is There a Lag?’, *International Economic Review*, Vol. 27(2): 265-83.

Hallonsten, O., Benner, M. and Holmberg, G. (2004). ‘Impacts of Large-Scale Research Facilities – A Socio-Economic Analysis’, *Research Policy Institute*, Lund University, Sweden.

Han, B. and Windsor, J. (2011). ‘User’s willingness to pay on social network sites’, *Journal of Computer Information Systems*, 51 (4): 31.

HM Treasury (2003). *The Green Book – Appraisal and Evaluation in Central Government*, Treasury Guidance, London: TSO.

HM Treasury (2006). *Stern Review: The Economics of Climate Change*, London.

Irvine, J. and Martin, B.R (1984). ‘CERN: past performance and future prospects. II. The scientific performance of the CERN accelerators’, *Research Policy*, 12: 247-284.

Jaffe, A.B. (1989). ‘Real effects of academic research’, *The American Economic Review*, 79(5): 957-970.

JASPERS (2013). *Project Preparation and CBA of RDI Infrastructure Project*, Staff Working Papers, JASPERS Knowledge Economy and Energy Division.

Johansson, P. O. (1991) *An introduction to Modern Welfare economics*, Cambridge, UK: Cambridge University Press.

Jura Consultants (2005). ‘Bolton’s Museum, Library and Archive services: An economic valuation’, report prepared for Bolton Metropolitan Borough Council and MLA North West, Edinburgh.

Kay, J.A., and Llewellyn Smith, C.H. (1985). ‘Science Policy and Public Spending’, *Fiscal Studies*, Vol. 6(3): 14-23.

Knight, F. (1921, 2nd edn 1964).*Risk, Uncertainty and Profit*, New York: Sentry Press.

Kulmus, G. (2012). ‘Preface’ in Kalmus, G., Brown, R., Evans, D., Gibson V. and Nickerson, R. (eds) *Physics at the high-energy frontier: the Large Hadron Collider project*, Philosophical Transactions of the Royal Society A, 370 (1961): 801-1004.

Lederman, L.M. (1984). ‘The Value of Fundamental Science’, *Scientific American*, 25: 34-41.

Maiani, L., with Bassoli R. (2012). *A caccia del bosone di Higgs. Magneti, governi, scienziati e particelle nell’impresa scientifica del secolo*, Mondadori.

Mansfield, E. (1991). ‘Academic Research and Industrial Innovation’, *Research Policy*, 20: 1-12.

Mansfield, E., J. Rapoport, A. Romeo, S. Wagner and G. Beardsley (1977). ‘Social and private rates of return from industrial innovations’, *Quarterly Journal of Economics*, Vol. 77: 221-240.

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

Mazzucato, M. (2011). *The Entrepreneurial State*, Demos, London, UK.

Mincer, J. (1974). *Schooling, experience and earnings*, New York: Columbia University Press.

Mouqué, D. (2012). ‘What are counterfactual impact evaluations teaching us about enterprise and innovation support?’, *Regional Focus*, 02/2012, DG Regional and Urban Policy, European Commission, Brussels.

Newman, M.E.J. (2004). ‘Who is the best connected scientist? A study of scientific coauthorship networks’, *Lecture Notes in Physics*, 650, 337-370.

NOAA (1993). ‘Report on the NOAA panel on contingent valuation’, *Federal Register*, vol. 58 n.10, Friday January 15, pp. 4602-4614.

O’Leary, Nigel C and Peter J. Sloane (2005). ‘The Return to a University Education in Great Britain’, *National Institute Economic Review*: 193: 75.
OECD (2003). *Turning Science into Business: Patenting and Licensing at Public Research Organisations*, Paris: OECD report.

OECD (2012). *Mortality Risk Valuation in Environment, Health and Transport Policies*, OECD Publishing.

OECD (2014). ‘Report on International Distributed Research Infrastructures’, report approved by the

OECD Committee for Scientific and Technological Policy.

Pancotti, C., Pellegrin, J., Vignetti, S. (2014). ‘Appraisal and selection of Research Infrastructures: Approaches, methods and practical implication’, draft paper produced in the frame of the research project “Cost/Benefit Analysis in the Research, Development and Innovation Sector”, part of the EIB University Research Sponsorship programme (EIBURS).

Pearce, D.W., Atkinson, G. and Mourato, S. (2006). Cost-Benefit Analysis and the Environment: Recent Developments, Paris: OECD Publishing.

Picot, A., Florio, M., Grove, N., Kranz, J. eds (forthcoming). *The economics of infrastructure provisioning*, MIT Press.

Pinski, G. and Narin, F. (1976). ‘Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics’, *Information processing and Management*, Vol. 12: 297-312.

Poincaré, H. (1908). *Science et méthode*, Paris: Flammarion.

Psacharopoulos G. (2009). *Returns to Investment in Higher Education. A European Survey*, Contribution to the Higher Education Funding Reform Project CHEPS-led consortium for the European Commission.

Psacharopoulos, G., and Harry A. Patrinos (2004). ‘Human Capital and Rates of Return’, in G. Jones and J. Jones (eds) *International Handbook of Economics of Education*, Cheltenham: Edward Elgar, pp. 1-57.

Romer P. M. (1990). ‘Endogenous technological change’, *Journal of Political Economy*, Vol. 98 (S),S71-S102.

Salter, A.J. and Martin, B.R. (2000). ‘The economic benefits of publicly funded basic research: a critical review’, *Research Policy*, 30: 509-532.

Samuelson, P.A. (1937). ‘A note on measurement of utility’, *The Review of Economic Studies*, 4(2): 155-161.

Scherer, F.M. (1965). ‘Corporate Inventive Output, Profits, and Growth’, *Journal of Political Economy*, Vol. 73(3): 290-97.

Schmied, H., (1975). *A study of economic utility resulting from CERN contracts*, CERN 75-5, CERN, Genève.

Schmookler, J. (1966). *Invention and economic growth*, Cambridge: Harvard U. Press.

Schopper, H. (2009). *The Lord of the Collider Rings at CERN, 1980-2000*, Springer, Berlin-Heidelberg.

Schultz, T.W. (1961). ‘Investment in Human Capital’, *American Economic Review* 51 (1) March: 1-17.

Science and Technology Facilities Council (2010). *New Light on Science. The Social & Economic Impact of the Daresbury Synchrotron Radiation Source, (1981 – 2008)*, The UK Astronomy Technology Centre, Royal Observatory.

Snowball, J.D. (2008). *Measuring the Value of Culture. Methods and Examples in Cultural Economics*, Springer.

SQW Consulting (2008). *Review of economic impacts relating to the location of large-scale science facilities in the UK*, Final report.

Starrett D.A. (1988). *Foundations of Public Economics*, Cambridge University Press.

Sund, B. (2010). *Economic evaluation, value of life, stated preference methodology and determinants of risks*, Örebro Studies in Economics 21, Örebro University.

Technopolis Group (2013). *Big Science and Innovation*, report prepared by Simmonds P., Kraemer-Mblua E., Horvath A., Stroyan J. and Zuijdam F., United Kingdom.

Times Higher Education (2011). *Citation averages, 2000-2010, by fields and year*. Published on line http://www.timeshighereducation.co.uk/415643.article.

Tohmo, T. (2004). 'Economic value of a local museum: Factors of willingness to pay', *Journal of socio-economics*, 33: 229-240.

Tuan, T. and Navrud, S. (2008). 'Capturing the benefits of preserving cultural heritage', *Journal of cultural heritage*, 9: 326-337.

Viscusi, W.K. and Aldy, J.E. (2003). 'The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World', *Journal of Risk and Uncertainty*, 27 (1): 5–76.

Vock, M., van Dolen, W., and de Ruyter, K. (2013). 'Understanding Willingness to Pay for Social Network Sites', *Journal of Service Research*, 16 (3): 311-325.

Vose, D. (2008). *Risk Analysis: A Quantitative Guide*, Great Britain: John Wiley and Sons.

Weikard, H-P. (2005). 'Why non-use values should not be used', Mansholt Graduate School Discussion Paper, 22, Wageningem University.

Weinberg, S. (1967). *Reflection on Big Science*, Cambridge, MA: MIT Press.

Westland, J.C. (2010). 'Critical mass and willingness to pay for social networks', *Electronic Commerce Research and Applications*, 9(1): 9-19.

Willis, K. (1994). 'Paying for heritage: what price for Durham cathedral?', *Journal of environmental planning and management*, 37(3): 267-278.

World Health Organization (2006). *Guidelines for conducting cost-benefit analysis of household energy and health interventions*, by Hutton G. and Rehfuess E., WHO Publication.

Annex 1. The theoretical model of CBA

In this Annex we briefly restate the CBA theory that we have in mind, which is at the basis of the model and notation to which we refer in the rest of the paper. An established model for social CBA is provided by Drèze and Stern (1987 and 1990).[76] Their model considers a benevolent social planner whose objective is to maximise social welfare by supplying goods according to a defined production plan. In this framework, a project is defined as the marginal change in the net supply of goods by the public sector. Solving the social planner's maximisation problem, subject to a scarcity constraint (and possibly other constraints, that we skip here), according to which private demand is equal to the public supply of goods, means to find the most socially deserving projects. Private demand is influenced by a set of variables, or 'signals', including prices, direct or indirect tax rates, rations on productions or consumptions, transfers. Hence, solving the social welfare maximisation problem is equivalent to determine the combination of signals (under the social planners' control) which has the most favourable social impact by changing the production plan. The environment is second best, meaning that the government cannot use optimal lump-sum and/or faces other policy constraints. The social planner's problem can be expressed as follows:

$$\begin{cases} Max\, V(s) \\ subject\ to \\ E(s) - z = 0 \end{cases} \tag{A1.1}$$

with $V(s)$ is the social welfare function to be maximised, which depends on a vector of signals s^{77} influencing private agents' behaviour; $E(s)$ is the unique environment in which the social planner operates and which determines private demand; z vector is the public production plan which has to fully satisfy private demand.

From the Lagrange function

$$L = V(s) - \lambda \cdot [E(s) - z], \tag{A1.2}$$

the First Order Condition to solve the constrained maximization problem is:

$$\frac{\partial L}{\partial s} = \frac{\partial V(s)}{\partial s} - \lambda \cdot \frac{\partial E(s)}{\partial s} = 0 \tag{A1.3}$$

where λ is the vector of Lagrange multipliers associated to the scarcity constraint. Lagrange multipliers are interpreted as the marginal effect of a change in the constraint upon the optimal value of the original objective function

$$\lambda = \frac{\partial V(s)}{\partial s} \cdot \left[\frac{\partial E(s)}{\partial s}\right]^{-1}. \tag{A1.4}$$

A project is defined as the marginal increase of public production dz. Such increase must be accommodated by the marginal change of private demand, thus it can be written:

$$dz = \frac{\partial E}{\partial s} ds. \tag{A1.5}$$

Plugging (A1.5) into (A1.4) shows that Lagrange multipliers are the first partial derivatives of the social welfare function around the optimum (first order condition) with respect to each good supplied by the social planner. In other terms, they reflect the social value of a marginal increase of the production plan (dz) on the increase of the social welfare function. Under some technical assumptions,

---

[76] See Florio (2014a) for a restatement of the theoretical CBA model.

[77] Not to be confused with the discount factor s_t mentioned in the rest of the paper.

Lagrange multipliers are exactly the (relative) prices assigned by the social planner to that marginal change of the production plan, or project, i.e. its shadow prices (v):

$$\lambda = v = \frac{dV(s)}{dz} \tag{A1.6}$$

i.e.

$$dV(s) = v \cdot dz. \tag{A1.7}$$

Market prices and shadow prices are often said by practitioners to be the same in perfectly competitive and socially efficient markets. In the reality, however, *all* markets in general (dis)equilibrium may be distorted by taxes, duties, subsidies, rigid exchange rates, rations on production or consumption, regulated tariffs, oligopoly or monopoly price setting and imperfect information. All these elements drive a wedge between the observed price and the marginal social value of resources. Unlike most of market prices in second-best economies, shadow prices reflect the social marginal value of a change of an output in the economy, i.e. the opportunity cost to the society of producing or consuming more or less of any good. As such, shadow prices should be used in order to evaluate the welfare impact of projects, rather than observed market prices.

Building on the Drèze-Stern theoretical framework, Cost-Benefit Analysis in a second best general equilibrium setting entails the assessment of whether an investment project dz makes profits at shadow prices v, so as to provide support for informed decision making. In the practice, a social CBA exercise consists of the calculation of the project economic net present value (NPV), defined as the difference between the discounted total social benefits and costs of the project occurring over a determined time horizon. On the benefit side, the project appraiser has to include the flow of operating revenues, valued at estimated shadow prices, positive externalities and other non-market positive effects, and the shadow residual value of the project; the latter is the present value at the last year of the analysis' time horizon of any net future benefit that the project would be able to generate because its economic life is not completely exhausted. On the cost side the analysis should include the capital cost of the investment and operating cost valued at shadow prices, plus negative externalities and other non-market negative effects.

The more general formula for the calculation of the economic net present value (NPV) of a project over a continuous infinite time is

$$NPV \; = \int_{t=0}^{\infty} e^{-\sigma t} \cdot f(t) dt \tag{A1.8}$$

where

$$f(t) = B(t) - C(t) \tag{A1.9}$$

is a function of benefits $B(t)$ and costs $C(t)$ over a continuous time, valued at shadow prices. This can also be written in a form similar to equation (A1.7):

$$f(t) = v(t) \cdot y(t) \tag{A1.10}$$

where $v(t)$ are shadow prices and $y(t)$ are quantities of project outputs net of inputs, occurring at time t.

The formula (A1.8) includes a discount factor $e^{-\sigma t}$ where σ is defined as

$$\sigma = ln(1 + r) \tag{A1.11}$$

and r is a social discount rate.

Discounting future inflows and outflows allows to aggregate costs and benefits, expressed in monetary terms, that occur in different periods of time. Due to impatience and a preference for present rather than future utility (a theme vastly discussed in the economic literature, see e.g. Samuelson, 1937; Caplin and Leahy, 2004; Boardman *et al.*, 2006), even when constant prices are used (so as to exclude an inflationary effects) the utility of spending or obtaining one Euro today is higher than one Euro tomorrow. Hence, in order to aggregate future benefits and costs and express them in terms of current money, they should be properly discounted at a social discount rate. The most common method of discounting is based on an exponential formula. Actually, from the definition of σ we can express the discount factor as:

$$e^{-\sigma t} = (1 + r)^{-t}. \tag{A1.12}$$

The NPV formula can be simplified considering a finite time horizon $\mathcal{T}$:

$$NPV \ = \int_{t=0}^{\mathcal{T}} e^{-\sigma t} \cdot f(t) dt \tag{A1.13}$$

In discrete and finite time, a practical approximation is:

$$NPV = \sum_{t=0}^{\mathcal{T}} v(t) \cdot y(t) \cdot (1 + r)^{-t} = \sum_{t=0}^{\mathcal{T}} \frac{B(t)}{(1+r)^t} - \sum_{t=0}^{\mathcal{T}} \frac{C(t)}{(1+r)^t} = PV_B - PV_C. \tag{A1.14}$$

with PV_B and PV_C are a simplified way to refer to the present value of benefits and costs respectively, discounted at rate r.

In this paper we do not consider the CBA theoretical issues related to non-marginal projects, those for which first-order conditions are insufficient for optimization. We assume that the kind of RIs we have in mind are large, but not large enough to change the features of the economy and of shadow prices within the time horizon of the evaluation (even if fundamental discoveries actually do that in the long run). CBA theory for large projects is discussed e.g. by Starrett (1988).

Annex 2. The model for valuing knowledge outputs

Elaborating on equation (5) and starting from the cost side, we have the present value of investment cost K, labour cost of scientific personnel L_s and other staff L_o, the present value of other operating cost O and negative externalities E.

In particular, the present value of labour costs can be defined as:

$$L_s = \sum_{t=0}^{\mathcal{T}} s_t \cdot l_{st} \tag{A2.1}$$

$$L_o = \sum_{t=0}^{\mathcal{T}} s_t \cdot l_{ot} \tag{A2.2}$$

where l_{st} and l_{ot} are the shadow wage of scientists and other staff respectively. To streamline the notation here we use $s_t = e^{-\sigma t}$. Similarly, we define the present value of capital, other operating costs and negative externalities as follows:

$$K = \sum_{t=0}^{\mathcal{T}} s_t \cdot k_t \tag{A2.3}$$

$$O = \sum_{t=0}^{\mathcal{T}} s_t \cdot o_t \tag{A2.4}$$

$$E = \sum_{t=0}^{\mathcal{T}} s_t \cdot \varepsilon_t \tag{A2.5}$$

where k_t, o_t and ε_t are respectively the annual capital costs, operating costs and value of negative externalities.

The total discounted value of the project cost is:

$$PV_{C_u} = \sum_{t=0}^{\mathcal{T}} s_t \cdot (k_t + l_{st} + l_{ot} + o_t + \varepsilon_t). \tag{A2.6}$$

As shown in Section 2, on the use-benefit side there is the sum of the present values of different effects. Focusing here on the value of knowledge output only, S, this is the sum of the present value of knowledge created in different times and by different scientists. The following notation is used:

- $t = 0,1,2, \dots \mathcal{T}$ indicates the time at which papers are produced.
- $i = 0,1,2 \dots, n$ indicates the various waves of paper production, where ‘0’ is used to refer to papers produced by RI’s insider scientists; ‘1’ to papers produced by other scientists citing papers of wave ‘0’; ‘2’ to papers produced by other scientists citing papers of wave ‘1’, etc.
- $P_{it} = p_{it} \cdot \zeta_{it}$ is the total social cost of producing papers of wave i at time t, computed as the number of papers of wave i produced at time t (p_{it}) multiplied by their shadow cost of production (ζ_{it}); for instance, P_{00} and P_{01} refer to the cost of papers written by scientists working at the RI in two different times, 0 and 1.
- $Q_{it} = q_{it} \cdot \xi_{it}$ is total social cost of citations received by papers of wave i from papers of wave $i+1$ at time t, computed as the number of quotations received (q_{it}), multiplied by the shadow cost of citations (ξ_{it}); for instance, q_{0t} is the number of citations received by papers belonging to wave 0 written by RI’s scientists at time t from paper belonging to wave 1. Note that the number of quotations received by papers of wave i is equivalent to the number of papers produced in wave $i+1$: $q_{it} = p_{i+1t}\ \forall i, t$.
- k_{it} is the number of references included in each paper p_{it}.
- $s_t = e^{-\sigma t} = (1+r)^{-t}$ is the discount factor, with r being the social discount rate that is assumed to be constant over time. Note that $s_0 = 1$ at time $t = 0$. The discounting process is used to progressively reduce the value of papers produced in more distant years, thus reflecting the rate of fall of the numéraire.

We assume that, if i-papers start to be written at time t, $i+1$ papers will start to be produced from time $t+1$, meaning that papers, once written, begin to receive citations starting from the next year. Then, considering a finite time horizon, the present value of knowledge output corresponding to *insider* papers is:

$$S_0 = \sum_{t=0}^{\mathcal{T}} s_t \cdot P_{0,t} + \sum_{t=1}^{\mathcal{T}} s_t \cdot Q_{0,t} =$$
$$= \left(P_{0,0} + s_1 \cdot P_{0,1} + s_2 \cdot P_{0,2} + \ldots\right) + \left(s_1 \cdot Q_{0,1} + s_2 \cdot Q_{0,2} + s_3 \cdot Q_{0,3} + \ldots\right). \qquad \text{(A2.7)}$$

Papers citing those produced by RI's scientists are, in turn, cited by other papers starting from time $t = 1$, but as mentioned above, the value imputable to the RI is the total cost of citations received by papers of waves 1, 2, etc., and the total cost of paper production divided by the number of references that papers contain. So, for papers of wave '1' we can write as follows:

$$S_1 = \sum_{t=1}^{\mathcal{T}} \frac{s_t \cdot P_{1t}}{k_{1t}} + \sum_{t=2}^{\mathcal{T}} s_t \cdot Q_{1t} = \left(\frac{s_1 \cdot P_{11}}{k_{11}} + \frac{s_2 \cdot P_{12}}{k_{12}} + \frac{s_3 \cdot P_{13}}{k_{13}} + \ldots\right) + \left(s_1 \cdot Q_{11} + s_2 \cdot Q_{12} + s_3 \cdot Q_{13} + \ldots\right). \text{ (A2.8)}$$

In general terms, the present value of knowledge output is then:

$$S = \sum_{t=0}^{\mathcal{T}} s_t \cdot P_{0t} + \sum_{i=1}^{n} \sum_{t=1}^{\mathcal{T}} \frac{s_t \cdot P_{it}}{k_{it}} + \sum_{i=0}^{n} \sum_{t=1}^{\mathcal{T}} s_t \cdot Q_{it}. \qquad \text{(A2.9)}$$

Simplifying the notation, we have that:

$$S = PV(P_0) + \left[PV\left(\frac{P_1}{k_1}\right) + PV\left(\frac{P_2}{k_2}\right) + \ldots + PV\left(\frac{P_n}{k_n}\right)\right] + \left[PV(Q_0) + PV(Q_1) + \ldots + PV(Q_n)\right]. \qquad \text{(A2.10)}$$

For the sake of simplicity, we could assume a common shadow wage w_s of 'insider' and 'outsider' scientists, a common shadow cost of papers ζ_{it} and citations ξ_{it} for every i and t, and a constant k_{it} for every i and t, meaning that every paper at any time contains the same (average) number of references. The number of citations over time, instead, is likely to be declining to reflect the obsolescence rate of knowledge.

Taking the cost and benefit side together (see equation (5) in the main text for the definition of other benefits) and considering now only the knowledge output benefit S, the net present value of a research infrastructure would result from:

$$NPV_{RI} = S - [K + L_s + L_o + O + E] + [T + H + C + A] + B_n = \qquad \text{(A2.11)}$$
$$= \left\{PV(P_0) + \left[PV\left(\frac{P_1}{k_1}\right) + PV\left(\frac{P_2}{k_2}\right) + \cdots + PV\left(\frac{P_n}{k_n}\right)\right] + [PV(Q_0) + PV(Q_1) + \cdots + PV(Q_n)]\right\} - [K + L_s + L_o + O + E] + [T + H + C + A]$$
$$+ B_n =$$
$$= \left[\sum_{t=0}^{\mathcal{T}} s_t \cdot p_{0t} \cdot \zeta_{0t} + \left(\sum_{i=1}^{n} \sum_{t=1}^{\mathcal{T}} \frac{s_t \cdot p_{it} \cdot \zeta_{it}}{k_{it}} + \sum_{i=0}^{n} \sum_{t=1}^{\mathcal{T}} s_t \cdot q_{it} \cdot \xi_{it}\right)\right] - \left[\sum_{t=0}^{\mathcal{T}} s_t \cdot k_t + \sum_{t=0}^{\mathcal{T}} s_t \cdot l_{st} + \sum_{t=0}^{\mathcal{T}} s_t \cdot l_{ot} +\right.$$
$$\left.\sum_{t=0}^{\mathcal{T}} s_t \cdot o_t + \sum_{t=0}^{\mathcal{T}} s_t \cdot \varepsilon_t\right] + [T + H + C + A] + B_n.$$

If the shadow cost of producing papers $\zeta_{i,t}$ is expressed by the value of scientific labour cost l_s, the social cost of the scientific personnel L_s cancels out with the social benefit of their knowledge output P_0 valued at marginal cost:

$$NPV_{RI} = \left[\left(\sum_{i=1}^{n} \sum_{t=1}^{\mathcal{T}} \frac{s_t \cdot p_{it} \cdot \zeta_{it}}{k_{it}} + \sum_{i=0}^{n} \sum_{t=1}^{\mathcal{T}} s_t \cdot q_{it} \cdot \xi_{it}\right)\right] - \left[\sum_{t=0}^{\mathcal{T}} s_t \cdot k_t + \sum_{t=0}^{\mathcal{T}} s_t \cdot l_{ot} + \sum_{t=0}^{\mathcal{T}} s_t \cdot o_t + \sum_{t=0}^{\mathcal{T}} s_t \cdot \varepsilon_t\right] +$$
$$[T + H + C + A] + B_n.$$
$$\text{(A2.12)}$$

Annex 3. Overview of the whole CBA model for research infrastructures

The NPV of research infrastructures over the time horizon $\mathcal{T}$ is defined as the difference between benefits and costs valued at shadow prices and discounted at the social discount rate r[78]. It can be decomposed in two parts: the net present value of use-benefits and costs NPV_u and the pure non-use value of discovery B_n.

$$NPV_{RI} = NPV_u + B_n = (PV_{B_u} - PV_{C_u}) + B_n. \tag{A3.1}$$

The present value of use-benefits PV_{B_u} is the sum of the economic value of knowledge output (S), technological externalities (T), human capital accumulation (H), cultural effects (C) and benefits of applied research to other users (A) . The present value of costs PV_{C_u} is the sum of the economic value of capital (K), labour cost of scientists (L_s) and other administrative and technical staff (L_o), other operating costs (O) and negative externalities if any (E). The net present value of unknown non-use benefits (B_n) refers to the possible effects of any discovery that the RI might find:

$$NPV_{RI} = [S + T + H + C + A] - [K + L_s + L_o + O + E] + B_n. \tag{A3.2}$$

The value of knowledge output is measured by the sum of the present value of papers signed by RI's scientists (P_{0t}), the value of subsequent flows of papers produced by other scientists that use or elaborate of the RI's scientists' results, divided by the number of references they contain ($\frac{P_{it}}{k_{it}}$, with $i = 1, \dots n$), and the value of citations each paper receives, as a proxy of the social recognition that the scientific community acknowledges to the paper ($Q_{it}$with $i = 0, \dots n$):

$$S = \sum_{t=0}^{\mathcal{T}} s_t \cdot P_{0t} + \sum_{i=1}^{n} \sum_{t=1}^{\mathcal{T}} \frac{s_t \cdot P_{it}}{k_{it}} + \sum_{i=0}^{n} \sum_{t=1}^{\mathcal{T}} s_t \cdot Q_{it}. \tag{A3.3}$$

The present value of technological spillovers is given by the discounted incremental social profits Π_{jt} generated by companies (j) of the RI's supply chain which have benefitted from a learning effect:

$$T = \sum_{j=1}^{J} \sum_{t=0}^{\mathcal{T}} s_t \cdot \Pi_{jt}. \tag{A3.4}$$

Human capital accumulation are valued as the increasing earnings (I) gained by RI's students and former employees (z), since the moment (φ) they leave the project, against a suitable counterfactual scenario:

$$H = \sum_{z=1}^{Z} \sum_{t=\varphi}^{\mathcal{T}} s_t \cdot I_{zt}. \tag{A3.5}$$

Outreach activities carried out by the RI produce cultural effects on the general public (g), which can be valued by estimating the willingness to pay of the general public W_{gt} for such activities:

$$C = \sum_{g=1}^{G} \sum_{t=1}^{\mathcal{T}} s_t \cdot W_{gt}. \tag{A3.6}$$

For the sake of completeness, we define the present value of benefits produced by applied research infrastructures on other users and the economic value of services provided by the RI (A) simply as:

$$A = \sum_{t=0}^{\mathcal{T}} s_t \cdot a_t \tag{A3.7}$$

The present value of operating costs can be expressed as:

---

[78] The discount factor is s_t.

$$PV_{C_u} = \sum_{t=0}^{\mathcal{T}} s_t \cdot (k_t + l_{st} + l_{ot} + o_t + \varepsilon_t), \tag{A3.8}$$

where k_t are annual capital costs, l_{st} and l_{ot} scientific labour and administrative/technical labour respectively, o_t other operating costs and ε_t the value of negative externalities.

If the marginal cost of scientists' labour cost is taken as a proxy of the value of knowledge outputs produced by scientists, l_{st} in equation (A3.8) and P_{0t} in equation (A3.3) cancel each other.

Finally, the residual value B_n captures two types of values related to the research discoveries: their quasi-option value (QOV_t)and the existence value (EXV_t):

$$B_n = QOV_t + EXV_t \tag{A3.9}$$

where, QOV_t is intrinsically uncertain and therefore not measurable, and simply assumed to be non-negative and then skipped; the existence value, on the other hand, can be proxied by stated or revealed willingness to pay for scientific research, and/or through benefit transfer.

In short, re-writing equation (5), the CBA model for pure and applied research infrastructures turns into the following equation:

$$NPV_{RI} = \left[\left(\sum_{i=1}^{n}\sum_{t=1}^{\mathcal{T}} \frac{s_t \cdot P_{it}}{k_{it}} + \sum_{i=0}^{n}\sum_{t=1}^{\mathcal{T}} s_t \cdot Q_{it}\right) + \left(\sum_{j=1}^{J}\sum_{t=0}^{\mathcal{T}} s_t \cdot \Pi_{jt}\right) + \left(\sum_{z=1}^{Z}\sum_{t=\varphi}^{\mathcal{T}} s_t \cdot I_{zt}\right) + \left(\sum_{g=1}^{G}\sum_{t=1}^{\mathcal{T}} s_t \cdot W_{gt}\right) + \left(\sum_{t=0}^{\mathcal{T}} s_t \cdot a_t\right)\right] - \left[\sum_{t=0}^{\mathcal{T}} s_t \cdot (k_t + l_{ot} + o_t + \varepsilon_t)\right] + (QOV_t + EXV_t). \tag{A3.10}$$

As B_n will usually be positive, the test is trivially passed for $NPV_u \geq 0$, while for $NPV_u < 0$, then $NPV_{RI} > 0$ if $EXV_t \geq -NPV_u$ and QOV_t is conservatively taken as zero.