



JRC SCIENCE FOR POLICY REPORT

Feasibility Study on Measuring the Causal Impact of Erasmus+ Actions

Sylke V. Schnepf
Corinna Ghirelli
Stefano Verzillo

2017



This publication is a Science for Policy report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

Contact information

Name: Sylke V. Schnepf

E-mail: sylke.schnepf@ec.europa.eu

JRC Science Hub

<https://ec.europa.eu/jrc>

JRC105692

EUR 28597 EN

PDF ISBN 978-92-79-68821-8 ISSN 1831-9424 doi:10.2760/262344

Luxembourg: Publications Office of the European Union, 2017

© European Union, 2017

Reuse is authorised provided the source is acknowledged. The reuse policy of European Commission documents is regulated by Decision 2011/833/EU (OJ L 330, 14.12.2011, p. 39).

For any use or reproduction of photos or other material that is not under the EU copyright, permission must be sought directly from the copyright holders.

How to cite this report: Schnepf, S. V., Ghirelli, C., Verzillo, S.; *Feasibility Study on Measuring the Causal Impact of Erasmus+ Actions*, EUR 28597 EN, Publications Office of the European Union, Luxembourg, 2017, ISBN 978-92-79-68821-8, doi:10.2760/262344, JRC105692

All images © European Union 2017, except: title page, source: <https://pixabay.com/en>

Causal impact evaluations of Erasmus+ actions is possible for preparations of the Erasmus+ post 2020 period

Abstract

Causal impact analysis of policies provides higher standards of measuring policy efficiency but requires the availability of suitable data. This feasibility report identifies sub-programmes of Erasmus+ which can be carried out to support post-2020 Erasmus+ developments.

Table of contents

Acknowledgements.....	1
Executive summary.....	2
1 Introduction.....	4
2 Why measuring the causal impact of Erasmus+?.....	6
2.1 Causal impact analysis for measuring policy effectiveness	6
2.2 How counterfactual impact evaluation works	6
2.3 Four main counterfactual impact evaluation methods.....	6
3 Data requirements for causal impact analysis	9
3.1 Administrative versus survey data	9
3.2 Preconditions of high quality data for causal impact analysis	9
3.2.1 Representative data	9
3.2.2 Information on participants and non-participants	10
3.2.3 Sample size.....	10
3.2.4 Outcome variables of interest	10
4 Possible short term CIE on higher education student mobility	12
4.1 Existing literature	12
4.2 Research gap	14
4.3 Data.....	15
4.3.1 Administrative data merged with survey data of graduates	16
4.3.2 Survey data of graduates	17
4.4 Methodology	19
4.5 Research questions	20
4.5.1 Cross-country comparison of causal impact of student mobility....	21
4.5.2 Country specific analysis on causal impact of student mobility	22
4.5.3 Policy objectives and research questions	23
4.6 Value added	23
4.7 Time frame.....	24
4.8 Risk assessment.....	24
5 Long term CIEs	25
5.1 Jean Monnet Chairs: research.....	25
5.1.1 Existing research	25
5.1.2 Data and outcome variables	26
5.1.3 Methodology	27
5.1.4 Value added.....	27
5.1.5 Time frame	27
5.1.6 Risk assessment	28

5.2	European Voluntary Service / European Solidarity Corpse	28
5.2.1	Existing research	29
5.2.2	Data	30
5.2.3	Outcome variables	30
5.2.4	Methodology	31
5.2.5	Value added	31
5.2.6	Time frame	31
5.2.7	Risk assessment	31
5.3	Pupil mobility	32
5.3.1	Existing research	32
5.3.2	Data	33
5.3.3	Outcome variables	35
5.3.4	Methodology	35
5.3.5	Value added	36
5.3.6	Time frame and risk assessment	36
5.4	Mobility of vocational trainees	37
5.4.1	Existing research	37
5.4.2	Data	37
5.4.3	Outcome variables	38
5.4.4	Methodology	38
5.4.5	Value added	39
5.4.6	Time frame	39
5.5	Master Loans	39
5.5.1	Existing research	40
5.5.2	Data	41
5.5.3	Outcome variables	41
5.5.4	Methodology	41
5.5.5	Value added	42
5.5.6	Time frame	42
5.5.7	Risk assessment	42
6	Erasmus+ actions difficult to evaluate with CIE	44
7	Erasmus+ actions impossible to measure with CIE	46
8	Conclusion	47
	References	48
	List of abbreviations and definitions	51

Acknowledgements

The authors thank Dalibor Mladenka (DG EAC), Sebastien Combeaud (DG EAC) and other participants for their comments and discussion in the meeting 'Impact evaluation of Erasmus+ actions' organised at DG EAC on 8 November. We thank Lene Mejer (DG EAC) for comments on an earlier draft. We are also grateful to Beatrice D'Hombres and Ian Vollbracht (JRC) for invaluable ideas and comments and Anna Shaleva (JRC) and Paola Biasi (JRC) for their contribution with background research.

Executive summary

Policy context

Erasmus+ is a major and emblematic EU policy. The probably most recognised element, higher student mobility celebrates its 30th anniversary this year and features an increase in student uptake from 3,200 to 272,500 participants from 1987 to 2014. The program aims among others to contribute to the achievements of the Europe 2020 Strategy objectives. It is without doubt that the program has also fostered many valuable cultural and academic links. In addition, there are a number of studies examining Erasmus+ and showing how positive the program is perceived by its participants.

Not much however is known about beneficial effects of Erasmus participation on labour market outcomes for participants and other areas of their lives. As part of the Better Regulation approach the Commission aims to evaluate the **causal impact of policies** to enable judgement about whether policies work and to draw policy conclusions for future planning. Counterfactual Impact Evaluation (CIE) is a core part of that approach and the most rigorous method in economics. However, CIE generally requires access to data that includes information on both those participating and not participating in the program. DGs however hold generally only data on programme participants. This is a major caveat that many DGs face for evaluating the causal impact of their policies.

Key conclusions

The complex data needs of CIE can best be met if data requirements are taken into account during the implementation of a policy. As far as Erasmus+ and many other EU policies are concerned, suitable data for CIE have not yet been collected in this systematic way for all beneficiary countries. Currently, this limits the scope of CIE analysis to data sets which were collected by other data holders.

This report discusses existing data sets that could be employed for CIE of Erasmus+ actions. The table below sets out the main areas where CIE approaches could be used and over what time horizon.

Main findings

The table presents only those objectives of Erasmus+ policies for which CIE could be used to estimate the policy's causal impact. Some objectives of Erasmus+ policies serve a wider non-quantifiable purpose of encouraging the study of, collaboration and interest in EU issues in the European and global academic community and can therefore not be evaluated directly.

In four Erasmus+ areas – i.e. European Voluntary Service (EVS)/ European Solidary Corps (ESC), pupil mobility, mobility of educational trainees and Master Loans - lack of suitable data means that CIE methods could only be applied once new data are collected via surveys. Such potential surveys could be coordinated with the data collection implementation of the 'Higher Education Knowledge Hub' created between DG EAC and the JRC.

For higher education mobility high quality data has either already been obtained or can be readily extracted from known sources. As a consequence, CIE can be conducted and results could be available to contribute to the preparations of the post 2020 Erasmus+ programme to reflect on how to best measure its impact. As regards higher education mobility, data obtained is representative for three countries: Italy, Germany and the UK. It should therefore be possible to conduct CIEs investigating the country specific impact of higher education mobility on labour market careers. This analysis, even though focused on a sub-set of the EU higher education systems, could then serve to inform wider policy considerations across the EU as a whole.

Related and future JRC work

In the future, the Competence Centre on Microeconomic Evaluation aims to conduct counterfactual impact evaluation on the policy areas discussed in this report. It is open to advice on the implementation of data collection during the policy implementation, counsel future decisions to be taken for investigating what policies work and conduct CIE of EU policies.

Table 1: Summary of possible counterfactual impact evaluations by Erasmus+ action

Action	Geographical coverage	Data	Data availability	Outcomes to be evaluated	Counterfactual impact evaluation method	Time frame
Higher education mobility	IT, UK, DE	admin + survey	obtained already	LM outcomes and job search attitudes	PSM, IV	1 year
EVS/ESC	EU	survey	could be collected	competences and attitudes	Diff-Diff + PSM	3 to 4 years
Pupils mobility	few countries (depending on data availability)	admin data merged with DG EAC data	data availability to be investigated for different countries separately	achievement and attainment	Diff-Diff + PSM	1 to 2 years for countries where merging possible
	EU	survey	to be collected	attitudes	Diff-Diff + PSM	2-3 years
Mobility of educational trainees	few countries (depending on data availability)	admin data merged with DG EAC data	data availability to be investigated for different countries separately	labour market outcomes	PSM, IV	1 to 2 years for countries where merging possible
	EU	survey	to be collected	attitudes	PSM	2 to 3 years
Master Loan	EU	survey	to be collected	LM outcomes and attitudes	RDD, Diff-Diff + PSM	Several years
JM Chairs	EU	admin + web data	to be extracted	Publications by JM Chairs and institute only	Diff Diff + PSM	1 year

Note: Diff-Diff refers to Difference in Differences approach. PSM denotes Propensity Score Matching. IV refers to Instrumental Variable approach. RDD is Regression Discontinuity Design. These are different methodologies applied for causal analysis. For an explanation see the text below.

1 Introduction

Policy context and aim of this report

Erasmus+ is a major and emblematic EU policy. It aims among others to contribute to the achievements of the Europe 2020 Strategy objectives (including a reduction of early school leavers and increase of tertiary education graduates) and the Strategic Framework for European Cooperation in Education and Training (ET 2020) which covers for example principles of lifelong learning and improvements of education quality and efficiency.

A variety of studies exist aiming to evaluate Erasmus+ and its predecessor programs. However, there is scope for improvement and sophistication of analyses. The predominant part of existing research does not focus on the causal link of Erasmus+ and its expected outcomes. Therefore, the Directorate-General Education, Youth, Sport and Culture (DG EAC) approached the Competence Centre on Microeconomic Evaluation (CC-ME) to produce a feasibility report discussing possibilities for evaluating the causal impact of Erasmus+. This work would enable contributions to preparations of post 2020 Erasmus+ programme and how to best measure its impact. This feasibility report responds to DG EAC's request.

In detail, this feasibility report answers the following questions: given the data requirements needed, how can the causal impact of Erasmus+ be evaluated? The report examines what could be done in the short and long term and which parts of Erasmus+ would be quite difficult to measure with counterfactual impact evaluation. The authors also provide a judgement on the value added of any counterfactual impact analysis for different Erasmus+ actions taking existing literature and policy characteristics (like duration, homogeneity of intervention and measures of objectives) into account.

What is causal impact evaluation and why?

Causal impact evaluation of policies enables judgement about how well policies work and drawing policy conclusions for future planning and budgetary decisions.

Counterfactual impact evaluation (CIE) is used to measure the causal impact of a policy. The distinctive feature of CIE is that it measures the impact of a given policy intervention by comparing the actual situation to the one that would have happened (the counterfactual) in the absence of the intervention. As such, CIE compares e.g. labour market outcomes of those who benefited from a policy (treated group) with outcomes of a so-called 'control group' of individuals, who are similar in all characteristics to the policy beneficiaries but their recipient status.

The emphasis on the evaluation of the causal impact of Erasmus+ is a direct consequence of the Better Regulation Guidelines. The Better Regulation Toolbox has recently been updated to include a stronger set of references to the importance of quantitative approaches and causal impact analysis. This includes, in particular, devoting more efforts to planning the collection of data during the policy definition phase with a view to enabling quantitative and qualitative evaluation. Last year's inauguration of the Joint Research Centre's Competence Centre on Microeconomic Evaluation (CC-ME) created a research group that aims entirely to support the European Commission's DGs by giving advice on data collection and conducting causal impact evaluations of EU policies.

Data collection during the policy implementation process

Causal impact evaluations require data on both the participants in the programme and non-participants. The data of non-participants provides the control group and therefore makes it possible to estimate the counterfactual: what would have happened if the participant had not taken up the programme. DGs generally collect data for programme participants only, since they benefit from funding. The collection of information on non-participants is more difficult, since they are not in any contractual relationship with DGs.

Nevertheless, data collection during policy implementation could be a solution for accessing these data. If data confidentiality is taken care of, data collection can be facilitated greatly since a considerable part of data is routinely generated during the policy implementation process. For example, the Erasmus+ action Master Loans was introduced in 2015 and is currently in its implementation phase. If banks distributing loans collect data on applicant students (both those who did and did not receive a student loan) counterfactual impact evaluation would be eased at low costs and possible within a small time frame.

In depth investigation on how to enlarge the coverage of existing DG EAC's administrative data sets is the key for making counterfactual impact evaluations and hence causal evaluations possible and feasible in the future. Any data to be used for analysis needs to be representative for the study population (including both "treated" and "untreated" individuals) and therefore meet basic survey design requirements.

What to do if administrative data was not collected during the policy implementation?

Nevertheless, with more and better access to administrative data sources from different data holders, the data requirements for counterfactual impact analysis can sometimes be met with existing data. These are survey or administrative data sets, at the national as well the regional levels, that were collected by organisations, agencies or data holders for purposes not related to measuring the impact of Erasmus+.

Assessment criteria for feasibility of causal impact measure

The judgements provided in the report of whether a counterfactual impact evaluation is feasible for preparations for Erasmus+ post 2020 depends crucially on the existence and access to suitable high quality representative data as well as on the features of the policy to be investigated. Programmes with heterogeneous and/or short-term activities are not suitable for counterfactual impact evaluation. Regarding the *long term outlook*, the authors consider mainly whether further representative data for programme participants and non-participants should be collected that fit the population addressed and objectives of the Erasmus+ action.

Structure of the report

The reminder is as follows: Section 2 describes in greater detail the advantages of counterfactual impact evaluation and explains its basic principles. Section 3 portrays data requirements to be met for conducting causal impact evaluation. Section 4 looks at short term and Section 5 at long term possibilities for measuring the causal impact of Erasmus+ actions. Sections 6 and 7 discuss areas that are more difficult to evaluate. Section 8 concludes.

2 Why measuring the causal impact of Erasmus+?

2.1 Causal impact analysis for measuring policy effectiveness

Over the 2014 to 2020 programming period, Erasmus+ will provide Euro 14.8 billion for supporting the development of actions, cooperation and tools linked to the objectives of the Europe 2020 strategy and its flagship initiatives in the areas of education and training, youth and sports. Article 21 of REGULATION (EU) No 1288/2013 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL and Article 21.2 of the Erasmus+ legal base set out the need to evaluate the impact of the programmes. In addition, the Better Regulation Agenda which has recently been updated stresses the importance of evaluations, mentioning the advantage of causal impact evaluation. The causal impact of a policy is important to judge whether policies work and to draw policy conclusions for future planning and budgetary decisions.

The Joint Research Centre's Competence Centre on Microeconomic Evaluation (CC-ME) aims at supporting the European Commission's DGs in conducting causal impact evaluations of their policies, in line with the growing focus on high quality evaluations where data availability allows. More and more DGs are currently signing up with CC-ME to measure the effectiveness of their policies causally.

2.2 How counterfactual impact evaluation works

Causal impact analyses employ counterfactual impact evaluation (CIE) methods. Its aim is to measure the causal or net effects of a policy, in our case Erasmus+ actions. In contrast to other types of evaluations like monitoring, progress evaluation and pure descriptive analyses, CIE aims at isolating the causal effect of a policy on its recipients. The distinctive feature of CIE is thus that it measures the impact of a given policy intervention by comparing the actual situation to the one that would have happened in the absence of the intervention. As such, CIE compares e.g. labour market outcomes of those who studied abroad funded by Erasmus+ (treated group) with outcomes of a so-called 'control group' of individuals, who are similar in all characteristics to the policy beneficiaries but their participation in the program.

By measuring the causal effect, counterfactual impact evaluation takes **selection bias** into account. Selection bias is the key problem hampering the examination of causal effects. For example, it is very likely that those students who decide to study abroad with Erasmus differ to those who decide not to be mobile. Mobile students are generally more motivated, have better language skills and are more confident than those students who do not intend to study abroad (Parey and Waldinger 2011, Rodrigues 2013). As such, students self-select themselves into becoming participants or non-participants depending on their characteristics. These characteristics, however, are also often associated with the outcome variable like i.e. income or language skills. If selection into participants and non-participants is not taken into account, the outcome measure can be biased and the estimated impact will not be a causal result of the policy. As discussed in the literature section below, most studies on Erasmus+ fail to consider the selection bias.

2.3 Four main counterfactual impact evaluation methods

How can selection bias be taken into account? There are four main methods which tackle the problem of selection in different ways.

Propensity score matching (PSM) creates a control group by matching individuals to the policy beneficiaries. For every individual in the treatment group (for example, for every Erasmus student studying abroad) another higher education student is identified in the data who shares the same characteristics except for the one of interest, in this case study abroad. PSM creates a score that summarises the characteristics used for

matching in one single number. Individuals are then matched based on this score. The causal effect is the difference in an outcome variable (like some measure of employability) between the treatment group (mobile students) and the matched control group (immobile students). Empirically, a problem of PSM is that individuals can only be matched on variables for which information is available in the data. Hence, PSM relies on the conditional independence assumption (CIA) which implies that besides the variables controlled for in the design, no other variables impact on selection of individuals into treatment and control groups. While PSM is relatively easy to apply if CIE has not been taken into account during the implementation phase of a policy, PSM needs a rich data set (many variables) to meet the CIA assumption.

The **differences-in-differences** (Diff-Diff) approach calculates the effect of a treatment (studying abroad) on an outcome (i.e. employability) by comparing the average change over time in the outcome variable for the treatment group to the average change over time for the control group. While the CIA assumption is not needed for Diff-Diff, it assumes that without treatment the average change between both groups would be parallel (so called parallel trend assumption). It is obvious that Diff-Diff needs data that provides at least two time points (panel/longitudinal data).

Regression discontinuity design (RDD) can be used if treatment (like participation in a mobility scheme) is determined by a cut-off point of a score. For example, if individuals' applications for a mobility programme receive a score and a threshold of this score determines on whether students will be selected into the programme, regression discontinuity design can be used. It is assumed that around the threshold the distribution of beneficiaries and non-beneficiaries is as good as random. Like Diff-Diff, RDD has specific data requirements: first, a cut-off point of a score variable needs to have been used for selecting into treatment; and second, the sample size around the cut-off point needs to be large. Fuzzy RDD allows for imprecise allocation of treatment around the cut-off point, which relates to a fourth method of CIE: instrumental variable approach.

For example, in Bologna universities select Erasmus participants from a pool of applicants by focusing on their university grades. Those applicants who reach a specific grade threshold will be selected to take part. In RDD, the causal effect of the programme is measured by comparing participants close to the threshold with non-participants close to the threshold. As such, RDD is close to an experiment, since it can be assumed that individuals close to both sides of the threshold are similar.

When participation in the programme is voluntary, the **instrumental variable** (IV) approach can be used to identify the impact of a programme. IV however needs an exogenous factor that explains very well the choice of participation in the programme. At the same time, this exogenous factor should not be related with the unobserved individual characteristics introducing the selection bias¹ such as motivation (this assumption is called exclusion restriction). For example, let the choice of participating in a training programme be determined by individual factors (e.g. motivation) and external factors (such as the distance between the home of the participant and the location where the programme is given). A valid instrument is an external factor (such as distance or exposure to the programme) that explains the choice of participation but is not correlated with individual's motivation. Hence, the instrumental variable approach allows identifying the causal effect of the programme on future labour market outcomes net of the selection bias. More details on IV (and other counterfactual impact evaluation methods) can be found in Angrist and Pischke (2009).

¹ The problem faced by the researcher in evaluating the causal effect of mobility programmes stems from the fact that the participation in the programme is voluntary, and hence likely to be associated with unobserved factors (such as motivation, or curiosity) that also affect the future labour market outcomes. This is the aforementioned selection bias.

Generally it is difficult to find a good instrumental variable. In the context of higher education mobility, a measure of students' exposure to mobility programmes could be used as a valid instrument to assess the impact of participation in mobility schemes on graduates' labour market outcomes. Students who are more exposed to mobility programmes are more likely to study abroad compared to students who are less exposed to these programmes. As is needed for an instrumental variable, the exposure to mobility programmes is unlikely to be correlated per se with individual motivation (and hence with better student's future labour market outcomes). This renders 'exposure to mobility programme' as a probably suitable instrumental variable.

The four quasi-experimental methods described above are generally employed after an intervention has taken place. Another possibility for measuring the causal effect is to use an experimental design during the implementation of a policy. DG EAC has commissioned some experimental studies which – to the knowledge of the authors – do not aim directly to measure the causal impact of the Erasmus+ actions described below². While experiments are often discussed as the 'gold standard' for measuring causal effects, they require considerable planning and need to satisfy ethical standards on treatment allocation.

The choice of a specific counterfactual impact evaluation will often depend on the richness of the data available. Data requirements for counterfactual impact evaluation are discussed in the following section.

² We are aware of two calls which aim at evaluations of DG EAC policies through experimental design. The call for proposals AECEA/04/2013 financed four projects on the prevention and reduction of early school leaving (CROCOOS, Jump@School, TITA, and SSP). A second call on 'Erasmus + Initiatives for Policy Innovation' aims to support transnational cooperation with a view to implement innovative policies under the leadership of high-level public authorities. The methodology entails randomised controlled trials if ethically acceptable. While these projects are of great interest and very innovative in their method choices, they do not directly aim to measure the causal impact of Erasmus+ actions on the population of beneficiaries.

3 Data requirements for causal impact analysis

The main reason for a lack of causal impact evaluation of Erasmus+ policies is the shortage of high quality data.

Recently, the 'Higher Education Knowledge Hub' has been created by DG EAC and the JRC. This Hub provides an opportunity for collecting high quality data needed for causal impact evaluation. Future surveys to be created and that are discussed in the following should be coordinated with the Higher Education Knowledge Hub if they are covered in its framework.

There are two possible data sources: administrative data and survey data.

3.1 Administrative versus survey data

Administrative data normally covers the entire population but generally only includes information on those characteristics important for administrative purposes. For example, administrative data from DG EAC or National Agencies only provides information on beneficiaries of Erasmus+. However, any causal analysis – as discussed above – also needs data on non-beneficiaries. Nevertheless, given that DG EAC has control over data collection, it has the possibility to determine on who and what information are collected in their administrative data for different actions of Erasmus+. In depth investigation on how to enlarge the coverage of existing DG EAC's administrative data sets is the key for making counterfactual impact evaluation and hence causal evaluation possible and feasible in the future. Negotiations at national level is key to implement data collection successfully.

Survey data refers to quantitative data that is collected from a study population using a questionnaire. Survey data needs to meet a variety of quality criteria which are discussed in survey design literature (i.e. Groves et al. 2009).

3.2 Preconditions of high quality data for causal impact analysis

There are several preconditions that high quality data need to fulfil for making causal impact analysis possible: representative data, data that includes information on participants and non-participants, big sample sizes and coverage of outcome variables of interest.

3.2.1 Representative data

High quality data is representative for the individuals examined. Administrative data is generally representative for the group of beneficiaries, since it covers the entire population. Survey data needs to be generated in a way that it is representative. For example, an examination of the impact of student mobility in French higher education institutes requires data that is representative for all French students in these institutes of a specific graduation year (so called study population). A representative sample can only be achieved by random sampling, which means that all French graduates of the study population need to have a non-zero chance of being in the sample. Random sampling can only be achieved if there is a list of all students in the study population (so called sampling frame) available, from which the sample can be drawn.³

³ This is a simplified discussion of sampling, since different kind of random sampling techniques can be used like stratification and clustering. If clustering is applied, a sampling frame is needed only for those clusters that are selected into the sample which facilitates sampling. For more information see the classical textbook of survey design by Groves et al. (2009).

Many techniques for collecting data, like snowball sampling and quota sampling or providing online links to a survey on a webpage do produce non-representative data.

If data is not representative, results cannot be generalised to the study population. The results are only valid for the sample, not for the population that is the focus of the examination. The size of the bias of the estimate cannot be determined. Standard errors (a measure of the accuracy of the estimates) cannot be estimated.

Besides random selection of survey participants, other quality features are important. High non-response of students jeopardizes the data quality. For example, email-surveys often feature a high non-response especially of those individuals who are not interested in the subject of the survey. Face-to-face surveys generally lead to higher response rates.

An extensive and comprehensive discussion of data generation for the analysis (data documentation) including information on how random sampling was conducted and possible biases due to non-response and other features is key and standard for quantitative research publications.

3.2.2 Information on participants and non-participants

Causal analysis depends on information on both, beneficiaries and non-beneficiaries of a policy. In the case of DG EAC administrative data provides information only on beneficiaries of a policy. However, if administrative data on a country's graduates were used, beneficiaries and non-beneficiaries of a policy would be covered, even though identifiers for beneficiaries might not be available. Survey data gains importance, if administrative data cannot be accessed. The survey can be administered to participants and non-participants.

3.2.3 Sample size

Administrative data covers the entire reference population.

In contrast, the costs of survey data are directly linked to its sample size, especially if face-to-face or telephone surveys are conducted (which are best in terms of minimising response bias). As a consequence, sample sizes are generally small especially once the focus is on sub-groups (for example students studying abroad). This generally impedes analysis that aims to disaggregate to small groups of interest (i.e. individuals studying abroad and coming from a disadvantaged family background). For example, the well-known cross-country surveys of international achievement 'Programme of International Student Assessment' (PISA) organised by the OECD allows for an average country sample size of about 7500⁴ pupils to analyse the educational achievement for different subgroups and its determinants. PISA sample sizes are certainly at the higher end of existing student surveys and allow focusing on tiny subgroups. For example, if Erasmus students account for 10 percent of graduates in a country, then we can examine only 100 Erasmus students if the sample size of a country covers 1000 students. With a sample of 100 students it is rather impossible to disaggregate further.⁵

3.2.4 Outcome variables of interest

Data for causal impact analysis needs to cover variables that measure the aim of the policy (i.e. mobility, employability or attitudes to Europe). If administrative data is used interesting outcome variables for the data analyst might not be covered, since these variables are not of interest for the data collector. Since DG EAC and National Agencies

⁴ See OECD (2014), which states that around 510,000 students participate in 65 countries in PISA 2012.

⁵ The standard error which needs to be calculated for every survey estimate provides a measure of precision of the estimate. The smaller the sample size the lower is precision around the estimate.

collect administrative data on Erasmus+ they are free to determine the variables covered.

If DG EAC conducts its own survey to analyse the impact of Erasmus+ it can determine the variables of interest collected and covered in the analysis. Data (like survey data) collected by the data user in order to answer research questions is called **primary data**. Collecting primary data is generally costly and time consuming given the need to plan, organise and implement a survey design and collect and edit the data for analysis. The advantage of collecting primary data is that all variables of interest, i.e. cultural values of Erasmus students, attitudinal changes etc., can be collected, since the data user has complete choice on which variables are collected.

In order to avoid costs of primary data collection, analysts tend to use survey data collected for other/many purposes and by other organisations. The advantage of such data is that it is normally readily available as well as its use endures only low costs compared to conducting a new specifically focussed survey. Its disadvantage is that important variables for the analysts' research might not be covered. Examples are the micro data available from general population surveys as labour force, household budget, income and living standard surveys which often include education variables. For Erasmus, some surveys will be discussed later in greater detail.

4 Possible short term CIE on higher education student mobility

This Section discusses a feasible counterfactual impact analysis that can be conducted in a shorter term. For higher education student mobility data can be made available and analysis can be conducted in time to be fed into the preparations for Erasmus+ post 2020.

The mobility of higher education students is part of Key Action 1 of Erasmus+. The main objective of students' mobility is to increase their skills, enhance their employability and improve their cultural awareness and intercultural skills. Student mobility should also contribute to personal and professional development and equip with transferable skills.

Higher education students generally choose between a study period of 3 to 12 months abroad at a partner's higher education institution or a traineeship work placement from 2 to 12 months abroad in an enterprise or any other relevant workplace.

Causal impact evaluation needs to take into account how individuals are selected into the treatment group. First, students self-select themselves by being aware of Erasmus, deciding to want to take part and writing an application. Higher education students receive a mobility grant if their application is selected by their higher education institute which needs to have been awarded the ERASMUS Charter for Higher Education (ECHE)⁶ by the responsible National Agency.

The 2017 annual work programme of DG EAC (DG EAC, 2016a) foresees that around 691 Million Euros will be spent on higher education mobility (including student and staff). This reflects almost one third of the entire Erasmus+ annual budget. Student mobility is therefore an important area within Erasmus+ to investigate. The year 2017 marks the 30th anniversary of higher education student mobility, which has increased markedly over the last decades showing the growing interest of students in the program and hence in experiencing different cultures, learning in foreign countries and improving language skills. The increased uptake and positive feedback by students (for example: Otero 2012; Engel 2010; Mitchel 2012) show the success and importance of the program.

There is also some research, examining the implications of students' mobility on a variety of characteristics. However, as will be outlined in the literature section below, not much is known about the causal impact of this Erasmus+ action on outcome variables measuring the main objectives of the program.

4.1 Existing literature

While there is considerable literature on higher education student mobility, Sorrenti (2015) states that literature often provides anecdotal and qualitative or quantitative evidence based on data that are not representative for graduates making therefore inferences to the population difficult.

⁶ ECHE is an accreditation granted by the European Commission giving the possibility to higher education institutions from Programme Countries to be eligible to apply and participate in learning and cooperation activities under ERASMUS+. The Charter outlines the fundamental principles an institution should adhere to in organising and implementing high quality mobility and cooperation and states the requisites it agrees to comply with in order to ensure high quality services and procedures as well as the provision of reliable and transparent information.

The predominant part of the literature on student mobility examines perceptions and characteristics of Erasmus students. Otero (2012) investigates results of a 2004/05 online survey from 30 countries on Erasmus students who self-report their perceptions. He finds that 58% of students reported that their Erasmus period had influenced their career-related attitudes to a large extent and 92% reported that the period abroad had changed their understanding of people from another cultural or ethnic background to at least some extent. Engel (2010) exploits a 2005/06 Erasmus student survey covering Eastern and Western European countries and conducted within the Valera⁷ Program. She shows that nearly 70 percent of former ERASMUS students are satisfied or very satisfied with their current work while only 11 per cent are dissatisfied. This literature on Erasmus students is interesting, because it provides information on the popularity of the program and the positive perceptions and high satisfaction of participants.

Some literature examines differences in a variety of outcomes comparing Erasmus and non-Erasmus students. Mitchel (2012) exploits a survey on university students in year 2011 in five countries: France, Germany, Italy, Spain and the UK. The data was collected only for one university per country. Random sampling was not applied so that the data is not representative, which – as discussed in Section 3 - makes it difficult to infer to the study population. Mitchel shows that compared to non-Erasmus students, students in the five universities who studied abroad report that they became more interested in other European countries and cultures and in the EU.

Of great interest is the Erasmus Impact Study (DG EAC, 2014a), which uses survey data on European countries to compare employability skills and personal traits between Erasmus and non-Erasmus beneficiaries in all countries together. A follow up Erasmus Impact Study (DG EAC, 2016d) extends the analysis by analysing the data at the regional level.⁸ The studies show that employability skills and personal traits are considerably higher for Erasmus compared to other students. Unemployment rates are lower for Erasmus students and cultural awareness of students are high. However, it is not quite clear whether these results are linked to the causal effect of the policy or individuals' selection.

A third smaller group of literature discusses this problem of selection bias. This group of literature approaches the measurement of impact differently to the studies discussed before by trying to assess the causal influence of the Erasmus action.

To the knowledge of the authors, the only available cross-national study (Rodrigues 2013) uses data from two projects funded by the European Framework Programmes: REFLEX (Research into Employment and Professional Flexibility) and HEGESCO (Higher Education as a Generator of Strategic Competences). Both projects consist of a large scale survey among graduates from higher education, 5 years after their graduation. The REFLEX project was carried out in 2005 in fourteen countries (Austria, Belgium's Flanders, Czech Republic, Estonia, Finland, France, Germany, Italy, Japan, the Netherlands, Norway, Portugal, Spain, United Kingdom), surveying around 70.000 graduates from ISCED 5A (International Standard Classification of Competences) programmes who got their degree in the academic year 1999/2000. HEGESCO, was conducted in five other European countries (Lithuania, Poland, Hungary, Slovenian and Turkey) in 2008 with a sample size of 30.000 graduates, finishing ISCED 5A⁹ programmes in the academic year 2002/2003. Rodrigues finds that student mobility is

⁷ For more information see Engel 2010.

⁸ A third study by the same contractor is on its way as this report is written.

⁹ In contrast to the more practically orientated ISCED 5B, tertiary ISCED 5A programmes (ISCED 5A) are largely theory-based and provide qualifications for entry to professions with high skill requirements.

associated with a significant increase in the probability to be mobile after graduation. Student mobility is also associated with a slightly longer time to find the first job after graduation. In addition, student mobility is associated with higher hourly earnings. In order to take selection bias into account, the counterfactual impact evaluation method PSM is applied. As discussed in Section 3, PSM is based on the conditional independence assumption. This means, effects found are only causal if the main variables for self-selection are covered in the data set (like motivation, personal skills, which are difficult to proxy).

The following other studies employ the counterfactual impact evaluation 'instrumental variable approach': Parey and Waldinger (2011) use German graduate data of the cohorts 1989, 1993, 1997, 2001 and 2005. The data is collected randomly and is therefore very likely to represent the population of German graduates. The authors find that studying abroad increases an individual's probability of working in a foreign country by about 15 percentage points. This is confirmed by Oosterbeeck and Webbink (2011) who show that the probability of living abroad is significantly increased for students studying abroad employing representative data from Dutch student cohorts of 1997 to 2002. Di Pietro (2013) exploits a nationally representative survey conducted by the Italian National Statistical Institute (ISTAT) in 2007 on graduates from Italian universities. He shows that studying abroad has a relatively large and statistically significant effect on the probability of being in employment 3 years after graduation. Sorrenti (2015) uses the Italian population data on graduates from AlmaLaurea but focuses on language proficiency. Study abroad programmes improve significantly language proficiency and the effect is highest for people from disadvantaged families in terms of socio-economic background.

4.2 Research gap

Given the very limited number of causal impact literature discussed above, there is general agreement that the big uptake of Erasmus+ student mobility during the last decades is far from being matched by causal evidence on possible related outcomes (Sorrenti, 2015; Di Pietro, 2013). Taking into account that in 2017 about 691 Million Euros will be spent on higher education mobility of staff and students (DG EAC 2016a), the existence of less than 10 studies actually measuring the causal impact and efficiency of the program is rather small.¹⁰

In addition, there are important gaps in knowledge which could guide the development of the policy:

- With the exception of Rodriguez (2013), causal impact evaluation studies and other literature focus only on a single country or an entire region. There is not much known about differences in the impact of Erasmus depending on the country setting. Higher education as well as labour market demands and attitudes of employers differ by country. As such it can be assumed that what is gained by studying abroad might differ by country settings. However, this is a completely ignored field of research.
- Student mobility in the framework of Erasmus has increased considerably during the last three decades of its existence. The last 30 years were also marked by globalisation, increasing international collaboration and the post-financial crisis. Within the changes of world economies, how has the causal impact of mobility

¹⁰ The European Commission's DG on Employment, Social Affairs and Inclusion identified a similar mismatch of causal impact evaluation studies and budget spent on European Social Fund Policies and therefore initiated pilot projects to measure the causal impact of ESF policies (Elia et al., 2015).

changed over time? The increased uptake of students could reflect that students experience an increased gain of mobility.

- Currently it is not known whether the gains of mobility differ by subject studied. The current Erasmus Impact studies provide information on fields studied by mobile students. However, it is not examined whether the mobility is more important for some than for other subjects. Students studying Science, Technology, Engineering and Maths (STEM) have been in the centre of attention recently. Do these students benefit in a different way from mobility than for example language students, who are the predominant part of mobile students?
- Sorrenti (2015) shows that language proficiency gains are highest for individuals from disadvantaged households. Di Pietro (2007) shows with Italian data that study abroad programs improve career prospects especially for individuals from disadvantaged backgrounds. This indicates that Erasmus mobility decreases social inequalities in labour market access and managerial professions, which would be a great success of the program. However, this has not been examined systematically up to now.
- Current literature does not examine whether mobility within and beyond Europe impact differently on labour market outcomes or labour mobility. This is of importance given the considerable Erasmus+ budget allocated to students studying outside Europe.
- Erasmus+ student mobility is just one mobility abroad students can choose from. In the UK student graduate cohort 2014/15, slightly more students went abroad in the framework of university arrangements than with Erasmus+. Nevertheless, given that we do not know much about the causal effect of Erasmus+, we cannot guess whether other mobility programmes have similar effects. And if the effect between Erasmus+ is different to that of other programmes, what could we learn from that?

In sum, causal impact analysis is needed to estimate the impact of Erasmus on skills, employability and personal development lines. Some recent studies using counterfactual impact evaluation have acknowledged this necessity. However, current literature is sparse and neither matches the importance of the Erasmus higher education mobility scheme in terms of number of students participating nor in terms of budget spent. General outcome factors investigated in existing studies regard whether students are more likely to live abroad and whether their employability has increased due to participation in Erasmus. However, existing studies generally do not differentiate the impact of Erasmus by socio-economic background, subject studied, whether study location was within or outside Europe and over time. In addition, studies are either focussing on regions or single countries, but do not compare results from different countries with varying education systems and labour market demands.

4.3 Data

Currently, no cross-national survey of European graduates is available that could be used for preparations of Erasmus+ post 2020. However, DG EAC's Life Long Learning Program funded a EUROGRAD feasibility study (Muehleck et al. 2016), which discusses the possibility of collecting European wide data on graduates in order to allow research on professional and personal developments of Europe's higher education graduates. A recently launched Council Recommendation outlines the need of a harmonised collection of cross-national longitudinal data on graduates. It is aimed to investigate whether important criteria for harmonising graduate data cross-nationally can be met and repeated to achieve longitudinal data.

This pilot offers a window of opportunity for future causal impact evaluation if the future graduate survey includes variables on participation in DG EAC education policies as well as a wide coverage of important background variables. Advice by counterfactual impact evaluation experts on survey design and variables to be covered could guarantee that future data for causal impact analysis is available.

However, with the focus on immediate data availability, this section discusses currently available secondary data meeting the minimum requirements for causal impact analysis of higher education mobility (discussed in Section 3).

Given that causal impact analysis of higher education students' mobility requires representative data sets on both, mobile and non-mobile, higher education students, administrative data by National Agencies cannot be used without being integrated with other sources, since it only covers in-depth information on mobile students.

It is unlikely, that educational administrative data alone could be used for the casual impact analysis. In practice, it is not reasonable to assume that administrative data on graduates include information on mobility during higher education enrolment. Administrative data either covers students before entering the labour market (educational data) or individuals having entered the labour market (labour market/social security data). The former by definition cannot include information on labour market outcome while the latter does not include information on student mobility.

As such, it is likely that administrative data of students or graduates would need to be linked to survey or additional administrative data. As such, any kind of educational administrative data would need to be integrated with additional data sources.

Another possibility would be to gain data from a survey of graduates (as is planned with the Eurograduate Survey).

The following describes currently available administrative data merged with survey data and survey data on graduates. The data sets refer to different graduate cohorts which faced different labour market entry challenges. However, given that all data sets cover several cohorts of graduate students, big variations between graduate cohorts possibly due to business cycles can be identified.

4.3.1 Administrative data merged with survey data of graduates

a) AlmaLaurea data for Italy

AlmaLaurea is a consortium of universities founded in 1994 which collects general information on Italian graduates and their first labour market experiences. The data covers information about careers of the population of students of 71 out of 88 Italian Universities in 2016. Students receive several individual questionnaires. First, students complete a questionnaire just before graduation. The response rate of students is around 90%. Second, graduates complete a questionnaire on individual labour market career after graduation one, three and five years after graduation. Response rates are between 80%-90% one year, around 80% three years and 75% five years after graduation. Given that population data are available, it will be possible to adjust for non-response bias to some degree. The data is collected annually.

The data covers the population of students providing therefore a sample size sufficiently big for disaggregating mobile students by socio-economic status and subject studied.

The survey covers Erasmus mobility and information on graduates' socio-economic status. Possible outcome variables are the following: employment status (after one,

three and five years), monthly net wage (after one, three and five years), type of contract (permanent vs fixed-term), and job satisfaction (after 5 years).

The data set can only be accessed and analysed in Bologna where the data are held.

b) Higher Education Statistics Agency (HESA) data merged with longitudinal graduate survey data for the UK

HESA data is population data of each student cohort in higher education institutes in the UK. Given that students have individual identifiers, different years of student cohort data can be merged into longitudinal data. In addition, HESA conducts two graduate surveys. First, the Destination of Leavers from Higher Education Records (DLHE) survey is administered to all graduates six months after graduation. It is carried out by higher education institutes and the data is returned to HESA. Second, every second year those graduates who took part in the first survey are contacted up to three and a half years later again by their higher education institute (DLHE longitudinal). Both surveys request information on early career developments after graduation.

HESA administrative data on student records can be merged with the DLHE data from both surveys. As such, the data include information on labour market careers.

HESA data could be attained for the graduation cohorts 2006/07, 2008/09, 2010/11, 2012/13 and 2014/15, which covers a sample of about 1 million graduates. With the exception of 2014/15 information from both DLHE surveys is available for all cohorts. The data includes information on whether a student studied abroad with Erasmus and in which country and socio-economic status of the student (social class and parental education). In terms of outcome variables the following are covered at both time points (6 months and 3 years after graduation): employment status with job title, country of employment, job characteristics for current / last job, standard occupational classification (SOC), employment status since graduation, employment sector, earnings and part-time/full-time employment. The HESA data set also includes school subject grades achieved before attending university. As such, as is very unusual for data sets on graduate students, information on ability before University enrolment can be used for the analysis.

Given the big sample size of the data, the impact of mobility can be examined for a variety of subgroups of students, for which up to now not much is known (see literature discussion above). For example, given the availability of cohort data over the last decade, changes in the impact of mobility can be examined. The causal impact of mobility can also be investigated by subject studied, socio-economic background and destination country.

4.3.2 Survey data of graduates

c) ISTAT survey on graduates (Inserimento professionale dei laureati) for Italy

Every three years, the Italian Statistical Office (ISTAT) conducts a representative survey on Italian graduates three years after graduation. This survey includes both, graduates who completed a three-year first degree course and a second-cycle university qualification. The survey includes information on Erasmus mobility, socio-economic background and a variety of other personal attributes.

Possible outcome variables to examine are employment status (employed, unemployed, not active), job characteristics like type of contract (open-ended and fixed-term), employment sector (private, public, non-profit), industry sector, size of the firm, part-time/full-time, number of hours worked, other job opportunities, earnings and

qualifications needed for the job. Furthermore, labour mobility (living currently abroad) as well as attitudes in job search (willing to move abroad, reservation wage, willingness to accept PT/FT jobs) can be examined.

It is possible to obtain data for graduation years 2004, 2007, and 2011. The sample size per year is increasing as shown below.

Graduation type	Year of investigation		
	2007	2011	2015
5-years degree (Laurea a ciclo unico)/pre-Bologna system	26,570	17,241	9,750
Master	n.a.	13,671	21,909
Bachelor	20,730	31,088	26,741
Total	47,300	62,000	58,400

d) HIS data on graduates in Germany

As for Germany, the survey is conducted by the Higher Education Information System (HIS) institute for every fourth graduation years since 1988/89. For the graduation cohort 2000/1 graduates were surveyed three times, 12 months, 5 years and 10 years after graduation. The graduation year 2012/13 received only one survey up to now in 2014. Data is currently prepared and can be obtained by February 2017.¹¹

While the survey population is surveyed randomly, non-response bias could be potentially high given that only 25 percent of graduates respond. Non-response weights are provided. Furthermore, the 2012/13 graduation data merges two surveys together which employed different survey instruments. This will need to be taken into account in the analysis (Fabian et al. 2016, pp. 44 -76).

The data contains information on students' socio-economic background and study characteristics including information on mobility.

Possible outcome variables are employment status (including job title), job characteristics for current and last job, time needed to find the job, number of applications, number of invitations to interviews, employment status since graduation, employment sector, size of firm, earnings, qualifications needed for the job, whether the company is also based abroad, flexibility during work, part time/full time and other factors like estimation of job opportunities given graduation, reasons for change of jobs and attitudes to lifelong learning.

The sample size is around 16,000 individuals in the graduation year 2012/13 thereby again allowing a disaggregation of the impact of student mobility by different factors like socio-economic background and field of study.

¹¹ Until cohort 2008/9 the question on whether a student studied abroad was only asked in the third survey, hence 10 years after graduation. As such, data after 2000/01 and before 2012/13 do not contain information on whether a student studied abroad. From 2012/13 the question on studying abroad was moved to the first survey (one year after graduation). As a consequence, the causal impact evaluation is restricted to the years 2000/01 and 2012/13.

4.4 Methodology

The data available and discussed in the previous section meets the data requirements needed for conducting causal impact evaluation by being representative for graduates in a country and including information on both participants and non-participants of Erasmus. In addition, the data has sufficiently high sample sizes allowing disaggregation of students by background of interest. Big sample sizes make it possible to estimate even small effects of a policy. All data sets discussed also include variables of interest for evaluating the policy (the latter will be discussed in the next Section in greater detail).

The available data cover only a small number of European countries. However, the authors are currently aiming to access additional administrative data for other countries which could be examined in the future (like Spain and Sweden). Currently, not much is known about country specific differences in the impact of student mobility. A country comparison will provide useful insights on differences in effects of Erasmus+ in 'typical sending' countries (such as Italy) versus 'typical receiving' countries (such as UK) and countries equally sending and receiving students (like Germany). In addition, even though focused on a sub-set of the EU higher education systems, the analysis could serve to inform wider policy considerations across the EU as a whole.

This high quality of the national data sets makes it possible to employ – depending on the research question – several counterfactual impact evaluation methods. Section 2 described counterfactual impact evaluation techniques briefly as well as the selection bias that needs to be taken into account in order to measure causal impact of student mobility.

In short, causal impact analysis takes into account that participations self-select themselves into the treatment group. Characteristics that determine self-selection however are often also related to outcome variables we want to measure. In the case of student mobility this can be shown as follows: students with better language skills, higher flexibility, higher motivation and positive attitudes to European values and different cultures are more likely to participate in student mobility programs than other students. If the effect of mobility programs is measured by just comparing mobile with non-mobile students' labour market chances, it is not taken into account that the both groups were already different before mobility took place. This selection bias impedes the evaluation of causal effects.

The following counterfactual impact evaluations can be employed given the available data described above:

a) PSM

As described in Section 2, PSM takes selection bias into account by matching beneficiaries and non-beneficiaries of the policy on characteristics associated with the outcome variables. The conditional independence assumption, however, needs to be met: the variables used for matching need to be entirely able to account for selection bias.

All data sets described above are very rich in terms of their background information of graduates. The UK and Italian data, very unusual for graduation data, also includes information on ability before University enrolment, which is probably a very good determinant of selection into mobile and non-mobile students.

Given the richness of the data, PSM seems a first method that can be used to determine the causal impact of mobility on a variety of outcome variables.

b) Diff-Diff

For using a differences-in-differences approach longitudinal data is needed that includes information on the outcome variable of interest before and after mobility abroad. These kinds of data are generally only available if surveys are collected for the aim of the analysis. As such, it is unlikely that Diff-Diff can be used for the analysis.

However, the future pilot Graduate survey, which will be conducted as a result of the Eurograduate feasibility study (Muehleck et al. 2016), could aim to measure student characteristics (i.e. 'attitudes to Europe') at time of University entry and after graduation. This would make it possible to use differences-in-differences to measure the causal effect of student mobility on attitudes to Europe.

c) RDD

RDD is very close to experimental design around the threshold of a continuous selection variable. It would be possible to use regression discontinuity design, if students were selected into the program on the basis of meeting a cut-off point of an application score. It is quite likely that some universities apply such a score to select eligible students. However this information is not shared. If in the future application scores were introduced and kept in a data base, regression discontinuity design could be easily implemented for causal impact evaluation.

d) IV

As described in Section 3, the instrumental data approach requires a variable, that is associated with selection into mobile and non-mobile student status but not associated with the outcome variable (for example labour market career). A good qualifying variable is the number of students taking up an Erasmus mobility grant by higher education institute and subject studied. This variable is available for UK data: for the UK, population data on graduates are available so that the variable can be computed based on the data set. For Germany and Italy, the instrumental variable clearly cannot be derived from the survey data provided. However, we assume that DG EAC could provide the data, which during the writing of this report is still under investigation. If we cannot receive the data from DG EAC, the different cohort data (from available surveys) can be used to estimate the instrumental variable.

In sum, for Germany, Italy and UK data PSM could be used in order to estimate the causal effect of mobility. In addition, for the UK and probably also Italy and Germany an instrumental variable approach could be used.

The choice of two different causal impact evaluation methods is of an advantage, since the robustness of results can be examined.

4.5 Research questions

Even though different in their coverage on information on graduates, all three data sets include variables which are similar. For these variables it is possible to conduct a counterfactual impact evaluation that compares the three countries. This is discussed in sub-section 4.1.4.1.

However, UK, German and Italian data have also very different features which allow for additional country specific analysis. This will be discussed in 4.1.4.2.

Section 4.1.4.3 relates the outlined research questions to the aim of the student mobility program, thereby showing their importance for measuring the efficiency of the program.

4.5.1 Cross-country comparison of causal impact of student mobility

In the following research questions are discussed that can be answered and compared between Germany, Italy and the UK given the data available. Such a cross-national analysis will compare two different graduation data sets: for Germany and the UK causal effects will be estimated 6 months after graduation. These results can then be compared to causal effects at 3 years after (UK and Italy) and 5 years after graduation (Germany).

As discussed in the literature section, up to now there does not exist any study that compares the following questions cross-nationally. As such, all questions listed below and analysed with a cross-national focus are a value added to existing knowledge on the impact of student mobility.

It is also very important to note, that given our large data sets the entire following questions can be answered by disaggregating students by low and high socio-economic background, Science, Technology, Engineering and Mathematics (STEM) subjects and non-STEM subjects studied and different cohorts. Results of the analysis could highlight whether Erasmus+ benefits different students in a different way.

i) How does mobility during higher education enrolment impact on graduates' **employment status**? Does the impact differ by socio-economic status, subject studied, country of graduation and graduation year?

If student mobility increases skills and employability, we would expect that graduates with experience abroad have an employment advantage to other graduates. The importance of the variable in the framework of Europe 2020 is obvious.

The causal analysis of employment status would gain if – as possible with the data – it is differentiated between full- and part-time employment taking gender differences into account. In addition, for Germany and Italy further disaggregation can be employed by investigating private and public employment and size of firm.

ii) How does mobility during higher education enrolment impact on **graduates' earnings**? Does the impact differ by socio-economic status, subject studied, country of graduation and graduation year?

It is widely acknowledged that better skills are associated with higher earnings (Becker 1964; Mincer 1958; Mincer 1974). We would assume to find an earning advantage of those graduates who took part in Erasmus. For short term stays abroad (3 to 6 months) it can be assumed that earning differentials are relatively small for students graduating from the same universities. However, given the big sample size especially for the UK, a significant effect can be demonstrated even for smaller earning advantages.

iii) How does mobility during higher education enrolment impact on individuals' **country of employment**? Does the impact differ by socio-economic status, subject studied, country of graduation and graduation year?

Cultural awareness and intercultural skills impact on graduates' choices and willingness to work abroad. It has been shown (Parey and Waldinger 2011) that Erasmus increases work placement abroad. Using the three countries data sets it will be possible to contribute to existing research and estimate, whether the causal effect differs between the three countries.

For Italy, it can also be examined whether working abroad matches the individuals' willingness to work abroad. For Germany, it can be examined whether there is a causal impact of student mobility on whether or not the graduate is working in multinational companies.

iv) How does mobility during higher education enrolment impact on **access to managerial professions**? Does the impact differ by socio-economic status, subject studied, country of graduation and graduation year?

Higher skills are also related to access to managerial professions. While questions on success in the labour market are administered in slightly different wordings in country surveys, access to higher level positions can still be compared between countries if these differences in survey design are taken into account.

For Italy and Germany, it could also be investigated whether the skills gained during higher education are needed for current employment. Mobile students could be more successful in matching their skills with their employment than other graduates, since they might be able to signal their skills better.

v) How does mobility during higher education enrolment impact differently **over time**? If there are changes over time, did students from lower socio-economic backgrounds gain more than their counterparts? Did students studying STEM subjects gain more than other students?

The uptake of the Erasmus student mobility program has considerably increased during the last decade. In addition, globalisation and international collaboration have equally gained more importance on the labour market. This raises the question, whether the increase of uptake of the mobility action might reflect increasing benefits of taking part in the program. Given that repeated cross-sectional data is available for all three countries, changes in the impact of student mobility can be examined.

4.5.2 Country specific analysis on causal impact of student mobility

For Italy the following research questions can be addressed:

vi) How does mobility during higher education enrolment impact on holding **permanent and fixed term contracts**? Does the impact differ by socio-economic status, subject studied, country of graduation and graduation year?

Being able to demonstrate higher skills is likely to be associated with better access to permanent contracts.

For Germany the following research questions can be addressed:

vii) How does mobility during higher education enrolment impact on **access to the first job**? Does the impact differ by socio-economic status, subject studied, country of graduation and graduation year?

The time it takes to find a job is an additional measure of employability. The German data includes many measures on access to first job, like number of applications, number of invitations to job interviews, methods of job search, time needed to find the first job.

viii) How does mobility during higher education enrolment impact on **attitudes to lifelong learning**? Does the impact differ by socio-economic status, subject studied, country of graduation and graduation year?

For the UK, the following research question can be addressed:

ix) Does the impact of student mobility differ to whether the student moved to a tertiary education institute **within or outside Europe**?

Since the share of students studying outside of Europe is small, for this question a disaggregation of the impact by socio-economic background and subject studied would not be possible.

4.5.3 Policy objectives and research questions

The main objectives of the action are the development of skills, increased employability, increased cultural awareness and intercultural skills.

The impact of student mobility on skills and employability can be addressed by measuring a variety of factors: employment status (question i), earnings (question ii), access to managerial professions (question iv), permanent vs fixed term contract (question vi) and access to the first job (question vii).

The impact of student mobility on increased cultural awareness and intercultural skills can be measured by individuals achieving to work in a different country than the one they graduated (question iii).

4.6 Value added

The value added of the analysis of the causal impact of Erasmus mobility is the following:

- The growing student uptake of higher education mobility paired with a large budget share allocated to this Erasmus+ for the action, needs to be met by research that shows that the program meets its objectives. This can only be shown with causal impact evaluation, hence the use of counterfactual impact evaluation. As the literature review shows (Section 4.1.2), literature examining the causal effect and hence the efficiency of the Erasmus student mobility is scarce. All the research questions discussed above can be examined with counterfactual impact evaluation.
- Even if research is taken into account that does not use causal impact evaluation, the research questions that can be examined with the available data address existing research gaps which help policy development (and which were developed in the literature review and discussed in Section 4.2). First, the entire analysis will compare the impact of student mobility across three countries and will thereby be – to the knowledge of the authors - the first to examine whether student mobility is equally successful in different countries. Second, there is very limited evidence showing that students from lower socio-economic background benefit more from student mobility. With the data available, it can be cross-nationally examined, whether students from lower socio-economic status benefit differently in terms of skills and employability. If benefits were higher for disadvantaged students, the mobility program would contribute to inequality reduction among graduates. Third, while students' interest in the Erasmus student mobility has constantly increased over the last decade, it is not known whether the impact of the program has increased over time too, as could be assumed given the growing importance of international cooperation. Fourth, to the knowledge of the authors none of the existing literature discusses whether the impact of student mobility differs by subject studied. This is of great importance, since if studying abroad is most important for some subject areas it would need to be further fostered and advertised. Fifth, the UK data is so large that analysis could also differentiate between students who studied in another European country or outside Europe.
- The research questions discussed manage to measure many dimensions of skills and employability, the two main objectives of the Erasmus+ action. In addition,

by being able to measure the country of employment, also an important dimension of intercultural skills can be examined.

4.7 Time frame

In the anticipation of the possibility to examine the research questions described above, the JRC has access to German, UK and the ISTAT Italian data set. These data sets could be exploited to estimate the causal impact of Erasmus+ on mainly labour market outcomes. Such an analysis could be conducted within several months. The time frame for causal impact evaluation based on other data sets like primary survey data, would take one to two years. Linking of existing administrative data with other data sources would also require more time.

4.8 Risk assessment

The risk associated with the execution of the above research and within the discussed time frame is small.

The following factors are outside of the control of JRC:

- Access to data on number of students taking up Erasmus student mobility for Germany and Italy, by higher education institute, subject studied and university. It would be very helpful to get access to these data which are stored by DG EAC.
- Access to the German data was postponed from November 2016 to February 2017. Data for the 2012/13 graduate cohort which was interviewed in 2014 is still in the process of being edited and cleaned by HIS.

5 Long term CIEs

5.1 Jean Monnet Chairs: research

The Jean Monnet (JM) Chairs action is part of JM Activities, which comprise six actions aiming at promoting excellence in teaching and research on EU studies around the world. Overall, the aim of JM Activities is to promote teaching, research and knowledge in the field of European Union studies. Among the JM activities, the JM Chairs is a type of programme that could be evaluated because it targets single professors and aims to enhance the beneficiaries' productivity in teaching but also in research activities.

A JM Chair is a teaching position with a specialisation in European Union studies for university professors. JM Chairs involve a flat rate financing system that grants 50,000 Euros for projects of 3 years. During the grant period, JM Chairs are also encouraged to publish their work, to participate and organize dissemination events with policy makers, organised civil society and schools, and to network with other JM Chairs or institutions.

At the institutional level, the objectives of JM Chairs are to increase the research and teaching capacity, and promote the production of knowledge on EU subjects through increased collaborations within and across institutions. Given these aims, appropriate outcome variables for evaluating this action would be focussed both on measuring teaching and research dimensions. However, it is important to note that the focus of this sub-section is on the outcome variable of the research dimension only, since data access and collection for and evaluation of the teaching dimension can be regarded as difficult. As a consequence, the reach of the policy evaluation discussed below is limited.

The 2017 annual work program of DG EAC (2016a) foresees a budget of 5.6 Million Euro for 134 JM Chairs. Although JM activities represent a small proportion of the overall Erasmus+ budget, the 2017 annual work programme acknowledges the strategic importance of JM, since it is the major programme at European level supporting studies on the European Integration process. According to the aforementioned document, additional activities to be performed by Jean Monnet Chairs will be encouraged (such as participation of young academics in teaching and research on European subjects), as well as the involvement of civil society in the activities. This indicates the need for a sound evaluation of the causal impact of the JM Chairs.

5.1.1 Existing research

To the knowledge of the authors, research evidence on the impact of JM Activities does not exist so far. A related and recent strand of the economic literature focuses on the flows of knowledge or uses indices based on scientific publications as individual/institutional outcome variables for a number of purposes. Bibliometric indices are used as proxy for both research productivity and research impact/quality to see how reforms of the university system in several European countries (e.g. Italy, France, Spain) affect the productivity of individual researchers or academic staff (see Zinovyeva and Bagues, 2015). Recent studies show how - at the departmental level - the average publication quality improves with average number of authors per paper, individual field diversity, number of published papers and foreign co-authors (Bosquet and Combes, 2013).

Current evaluations of Erasmus+ aims at assessing the impact of JM activities by means of qualitative analysis using case studies, interviews (on a selection of JM grant holders as well as of non-beneficiary organisations which deliver teaching and research on the EU), descriptive statistics based on a survey of JM grant holders, and an analysis of a

sample of 50 final reports of Jean Monnet grants. A quantitative causal impact evaluation of the research dimension of JM Chairs would very well complement qualitative studies.

As far as the research dimension is concerned, the research gap on the impact of JM Chairs can be addressed by considering researcher's productivity measures as individual outcomes of interest.

Available research measures to be used as a proxy for the production of knowledge could be for instance the number of publications in EU related matters at the individual and/or institution level by field of study (that might be by department or scientific discipline).

5.1.2 Data and outcome variables

As already discussed, funds of JM Chairs should foster teaching, research, and the production of knowledge on EU related matters. Hence, relevant outcomes to assess the impact of JM Chairs are teaching related outcomes (e.g. student satisfaction) and research measures such as publication indices on EU related studies. Given the fact that teaching outcomes are usually not available due to the non-comparability of self-reported student surveys or questionnaires across institutions, research indices are the only feasible measures to be collected for this study. The rationale for using these indices is that publication measures on EU related studies might be considered as good proxies of the production and quality of research on these topics.

For instance, one could construct publication indices for researchers belonging to target institutions focusing on the number of publications on EU related matters. Based on this index one could compare whether bibliometric indicators of researchers who received a JM Chair and/or are affiliated to institutions that host a JM Chair (maybe considering positive externalities of JM Chair's beneficiaries) grow faster than those of other similar researchers. More precisely, the following measures could be considered: number of publications as a proxy of scientific productivity, number citations as a proxy for research impact/quality, and number of co-authors as a measure of the research network. These outcomes could be extracted from the online databases on researchers' publications worldwide, i.e. Scopus, Google Scholar or Web of Science. These databases collect abstracts and citations of papers published by peer-reviewed scientific journals, of books and conference proceedings for different fields. For each publication, they report the main affiliation of the author(s), which can be used also to associate the publications to different institutions and hence construct appropriate indices. The information on the participation of different institutions and on beneficiaries of the Jean Monnet Chairs can be downloaded from the JM directory of the EACEA website which contains information of all JM Chairs granted from 1990 (when the JM activities were launched) onwards.

The time frame needed for the analysis would depend on the years for which the names and affiliations of participants in JM Chairs actions will be available from DG EAC, and on the quality of the information extracted from the internet for past publications. As conservative measure, a good starting point would be already to focus on the period since 2000.

The analysis would be conducted at both the individual and the institutional levels according to the field of study. The counterfactual group could be created by matching individuals and institutes on their institutional and bibliographic characteristics. Therefore, one would need to collect information about the quality of EU higher education institutions and their members (e.g. based on the European Tertiary Education Register (ETER) or bibliometric indices), so as to compare outcomes across similar researchers and their institutions. Some of this information would be extracted from the web.

5.1.3 Methodology

Since funding for Chairs is likely to impact on both the single researcher as well as the entire team of a research institute, the analysis would be conducted at the individual and institutional level by field of study. The causal impact of JM Chairs on the research production of the department hosting the Chairs could be evaluated by means of matching and Diff-Diff approach (see a description of these methods in Section 2). That is, the impact of receiving JM Chairs could be assessed by measuring indicators of research quality and production on EU related matters for all members belonging to EU institutions, and then by comparing average changes in pre- and post-funding publication indices for individuals and institutions receiving the JM Chairs (treated group) with similar ones that do not receive the JM Chairs (control group). The comparability between treated and control units (institutions or individuals) will be enhanced by means of matching techniques at the appropriate level.

Note that the matching exercise requires information about relevant characteristics of the units of analysis. The matching at the institutional level would require characteristics of the institutions (among which the discipline, past bibliometric indices and information from ETER could be employed as proxies for quality of the institutions). Geographically, the analysis at the institutional level could cover the EU member states, so as to enhance the comparability of institutions from different countries. However, it could also be restricted to a sub-set of countries where the higher education environments are homogeneous across institutions. By contrast, matching at the individual level would require matching on individual characteristics. In addition to past bibliometric indices (that can be also extracted from the web) it would be important also to match on individual characteristics such as research field, gender, age, and seniority. Such information cannot be retrieved from the web for all European countries. This makes a cross-national individual level analysis more problematic. One possibility would be to restrict the individual level analysis to a small number countries for which access to the names and characteristics of the population of the academic staff could be accessed. For instance, for Italy and the UK data sets on all academic staff in the country are available. The database for Italy is collected by the Italian Ministry of Education and managed by Cineca. It is public. For the UK, similar information is provided by the UK Higher Education Statistical Agency (HESA). Individual characteristics of professors would definitely enhance the quality of matching and therefore of the analysis.

5.1.4 Value added

Up to now, a quantitative analysis of the causal impact of JM Chairs on research activities has not been conducted. As a consequence, an analysis as described above would very well complement current qualitative analysis of the programme.

While the focus of the causal impact analysis would be limited to the research dimension of the programme, publications indices are interesting outcomes because they provide quantitative measures of flows and quality of knowledge/research and hence allow assessing the impact of JM Chairs based on quantitative techniques. The value added of this analysis is twofold. The approach would also be innovative and cost-efficient since it exploits data already available on the web to build up quantitative measure of research production and quality.

5.1.5 Time frame

The analysis could be organized in the following way: (i) extraction of the data; (ii) construction of the publication indices; (iii) counterfactual analysis and (iv) interpretation and dissemination of the results.

Phase (i) on the data extraction would be a crucial part of the analysis. This would entail to develop programming techniques. For the individual level analysis, it would be

necessary to get access to data on individual characteristics of academic staff at least in Italy and UK. While Italian data are public, getting access to the UK data may take time. Phase (ii) would allow constructing the following bibliometric variables: (i) outcome variables to measure the performance of individuals and institutions in terms of research and knowledge production; (ii) measures of research quality that would be needed for the matching approach. This phase is pivotal, since it requires both the design and the implementation of indices. It is likely that this phase requires about three months. Once the database would be finalised, the analysis would be straightforward. Hence, 6 weeks would be sufficient to run the difference in difference approach combined with matching (Phase iii). Interpretation and dissemination (Phase iv) would require one to two months.

5.1.6 Risk assessment

Due to the large amount of publication data needed for the analysis, the extraction phase will be time-consuming. Hence, a small risk of delay exists in the data extraction phase due to the need of checking that the extraction procedure is correct.

5.2 European Voluntary Service / European Solidarity Corps

For the future, the European Voluntary Service (EVS) and from 2018 the European Solidarity Corps (ESC) could be suitable for a counterfactual impact evaluation. EVS allows young people aged 17-30 to engage in unpaid and full-time voluntary activity in a host organisation within or outside the EU.¹² The volunteer service involves many different kinds of activities in various areas (e.g. youth work, cultural activities, social care, environmental protection, civic engagement, development cooperation, non-formal education programmes). Volunteers can engage in short-term or long-term service. Short-term EVS last between 2 weeks and 2 months, so that effects might be small and as a consequence more difficult to be identified as statistically significant (in contrast to being due to sampling variation). By contrast, in long-term EVS the individual service can last between 2 months and one year, which is likely to render bigger sized effects. As part of Key Action 1 of Erasmus+, the individuals are the action's beneficiaries.

Organisations which need to hold an EVS accreditation are in charge of managing the service for the beneficiaries. In addition, they apply to the call for expression of interest proposing an EVS project. An EVS project can include just one to up to 30 volunteers. During the application process, the presented EVS projects are assessed against the relevance of the project, the quality of its design and implementation, impact and dissemination. The project itself can last from 3 to 24 months.¹³ Once a project is awarded, eligible youths can apply directly through the participating organisations.

The objectives of EVS are to reach out to marginalised young people and promote the following principles: diversity, intercultural and inter-religious dialogue, common values of freedom, tolerance and respect of human rights literacy, critical thinking and sense of initiative of young people. Appropriate outcomes for evaluating the impact of this action

¹² Travel expenses, accommodation, food, local transportation, health insurance, language lessons and monthly allowance are covered and financed by the European Commission (the volunteer contributes only a 10% of the travel expenses).

¹³ EVS can be also involved in large-scale EVS events and strategic EVS projects. Large-scale EVS events entail at least 30 volunteers and must include complementary activities (i.e. conferences, seminars, meetings and workshops) in addition to the EVS activities. Strategic EVS projects are granted to experienced EVS coordinating organisations for projects aiming at generating solid and systemic impact. They implement standard EVS activities involving several volunteers and may include complementary activities.

could be measures of tolerance and attachment to European values. It is unlikely to find measures for the objectives in administrative data. As a consequence, measures would need to be collected through surveys.

According to the 2017 annual work program of DG EAC (2016a), 3 Million Euros are allocated to the EVS insurance,¹⁴ while 400,000 Euros are allocated to large-scale EVS events. The EVS budget has been increased in 2017 so as to launch the implementation of the ESC with the aim to further strengthening the impact of the action, as well as to open ESC to an even broader range of young people and organisations (DG EAC, 2016a p. 16). This witnesses the growing interest for EVS and calls for the need of evaluating the action through counterfactual impact methods.

5.2.1 Existing research

Since volunteering entails the provision of services for free, its value is not captured by official GDP statistics. According to an article from 'The Economist' the economic value of volunteering is often ignored by statisticians but can be estimated with data at hands and under some assumptions.¹⁵ In addition, the article claims that volunteering produces also private benefits (in terms of improving health and employment prospects of volunteers) and social positive externalities (e.g. decreasing crime activity). Below we review a number of scientific studies providing some empirical evidence in favour of these claims. Descriptive evidence is based on self-assessments collected through surveys on volunteers and voluntary organisations on the impact of volunteering. For a review of this literature, see Huiting Wu (2011). In addition, few studies provide causal evidence of the impact of volunteering on the volunteers' well-being, or on economic development and social inclusion.

A study by Corporation for National and Community Service (2008) evaluates the impact of AmeriCorps volunteering programme on the career prospects, civic engagement and well-being of volunteers. The study is based on longitudinal data and implements a Diff-Diff approach by comparing a group of individuals who participated in AmeriCorps in 1999-2000 with a similar group of individuals who expressed interest in joining AmeriCorps but did not enrol. The results show that, due to participation in the programme, AmeriCorps members are more likely to pursue public service career. In addition, results suggest positive long-lasting effects on civic engagement, and life satisfaction (measured 8 years after completion of voluntary service). Fujiwara and Kawachi (2008) study the causal impact of volunteer activity and community participation (as proxy for social capital) on physical and mental health in the U.S. This study is cross-sectional. In order to disentangle causal effect of participation in volunteering activity from unobservable individual characteristics such as personality or early childhood environment, the authors compare adult twins with different levels of social capital (proxied among others by volunteering). The causal effect is identified under the assumption that twins are similar in terms of personality and family background so that any difference in the outcomes is due to social capital. They find evidence of positive causal effect of social capital on physical health. In addition, the positive association between social capital on health is discussed in Kawachi and Berkman (2000).

Another strand of literature provides evidence in favour of the negative association between social capital and crime rates (e.g. Kawaguki et al. 1999). These associations are not necessarily causal. It is notable that the discussed evidence is restricted to the

¹⁴ The EVS insurance scheme aims at covering the risks run by the volunteers taking part in the EVS.

¹⁵ See 'The Economist' (12th September 2014). The economic value of volunteering can be estimated by calculating how much it cost to society to get these voluntary services produced if the volunteers had to be paid under regular contracts.

US. To the knowledge of the authors, causal impact of volunteering has not been investigated in the European context yet.

While the evidence discussed above refers to volunteering in general, to the knowledge of the authors the causal impact of EVS has not been examined so far. The EVS action has been evaluated only by participants' self-assessments of the programme using surveys. Based on the answers of the respondents, EVS has a positive influence on volunteers' lives. They gain self-confidence, independence, increase their capability of dealing with life difficulties and cultural awareness (e.g. Structure for Operational Support, 1999; DG EAC, 2016b).

In contrast to participants' assessment of the EVS, the ongoing EAC evaluation of Erasmus+ aims at evaluating the causal impact of EVS by means of counterfactual impact evaluation. To do this, the tender of this evaluation of Erasmus+ is intending to draw a sample of participating and non-participating young people in EVS.¹⁶ The tender foresees running the survey twice, so to measure the outcomes of interest before and after participation. This would allow for conducting a Diff-Diff approach in order to evaluate short-term EVS. Since short-term effects – as discussed above – are likely to be rather small and difficult to find to be significant, another approach could be to evaluate the causal impact of long-term EVS.

5.2.2 Data

To the knowledge of the authors, survey data containing measures of the EVS objectives are not available for a big enough sample of European youths. Even if this were the case, the survey should contain information on young people who participated in EVS mobility. Therefore, DG EAC would need to conduct its own survey in order to analyse the causal impact of EVS on a number of attitudes of beneficiaries. An advantage of this approach is DG EAC's choice of variables of interest for the analysis. The disadvantage is that collecting data is generally costly. It needs time to plan, organise and implement a survey design and collect and edit the data for analysis.

The duration of EVS can take up to 12 months, as such its impact on competences and career prospects might be more sizable than those found for youth exchanges.

In terms of survey design, it is feasible to first randomly draw a sample of organisations from the population of participating EVS organisations. Selected organisations would need to provide a list of all EVS applicants. Using this list a random sample of applicants (participants and non-participants) would be selected in each EVS organization of the sample. Because of the cluster design, the sample size of applicants needs to be increased compared to random sampling. It is important that once the sample is randomly selected, the survey to be conducted would need to include in-depth information on ability levels and socio-economic background to take selection bias into account.

5.2.3 Outcome variables

Given the major objectives of the EVS action, the following outcomes could be considered:

- (i) measures of competences: e.g. competences in communication skills and languages, digital competences;

¹⁶ Currently there is no documentation available discussing how random samples of young people are secured (see Section 3) and clustering of individuals in organizations is taken into account for the calculation of the sample size.

- (ii) attitudes: e.g. towards European values and cultural diversity, degree of tolerance and curiosity, participation in social and political issues, solidarity, increasing concern about inequalities and
- (iii) measures of job search attitudes, career aspirations, interest in achieving higher education, entrepreneurial skills.

5.2.4 Methodology

Two alternative methodologies could be used in order to measure the impact of the action on the aforementioned outcomes. If pre- and post-survey data is available, a Diff-Diff approach could be employed. The Diff-Diff approach would allow comparing the average change in relevant outcomes (i.e. interest in European values) before and after participation between participants and non-participants. The corresponding average change for non-participants represents the baseline evolution of the outcomes – that is it represents how the outcomes of participants would have evolved over time if the individuals had not participated in the EVS (the counterfactual). This requires that non-participants should be as similar as possible to participants. To ensure this, it is important that the questionnaire reports information on the socio-economic and educational background of applicants. (For more information see Section 2.)

If instead only post-survey data is available, the analysis could rely on PSM. The causal effect of participating in EVS would then be identified by comparing outcomes between participants and non-participants matched on relevant background characteristics that are likely to drive the participation choice. The quality of the approach depends on the richness of the survey data on individuals' background information.

5.2.5 Value added

Evaluations of the impact of EVS/ESC are drawn from descriptive evidence while a causal impact has not been shown yet. It is therefore not known whether EVS works and achieves its objectives. Given the growing emphasis attributed to EVS within Erasmus+ and the budget allocated, the importance of causal analysis on EVS has grown too. Therefore, a sound and robust analysis on the causal impact of EVS is needed to complement existing self-assessments of future ESC participants. In addition, the empirical analysis proposed in this chapter would represent added value to the existing literature on the impact of volunteering service, complementing the existing evidence on the causal impact of US volunteering programmes.

5.2.6 Time frame

Given the lack of individual level data on EVS (and future ESC) participants and non-participants, the research project illustrated above cannot feed into preparations for Erasmus+ post 2020. By contrast, the aforementioned analysis requires collecting individual data beforehand, which in turn requires time and resources for the planning and the implementation of the survey design. However, results could be provided in time for the final evaluation of Erasmus+ due by 30 June 2022. If this is of interest for DG EAC, the data collection should be planned as soon as possible.

The data collection should last at least 3 years in order to measure the outcomes of interest at most 2 years after the completion of the EVS / ESC service. Given the importance of the quality of the data for the final results, considerable amount of time and resources should be invested into the planning and design of the survey (at least 6 months).

5.2.7 Risk assessment

The risks associated with the above research project are mostly associated with the difficulties of ensuring high quality data collection through surveys. It is important that

selected EVS organisations can provide a list of EVS applicants that include contact details.

In particular, the JRC will take care about the following issues, so as to ensure the quality of the collected data:

- Decide the country coverage by maximising the representativeness of the country-specific samples with respect to their population
- Minimise low-response rate
- Design the questionnaires so as to maximise high quality responses.

Once the data collection procedure is sound and can be carried out properly, there are no risks associated to the analysis.

5.3 Pupil mobility

With more than 103,500 participants in 2014, this action represents a substantial area of Erasmus+ that accounted for an overall spending of 124 million euro (DG EAC, 2015b). Although mobility of pupils is not a mobility project *per se* but rather an activity embedded in strategic partnerships (Key Action 2), the fact that it targets pupils makes it very interesting for an impact evaluation: mobility at the early stage of the educational pathway may potentially contribute to lessening educational inequalities set during the early years and thereby improving children's educational outcomes and future opportunities.

Pupil mobility is organized by schools in the framework of Strategic Partnerships, which aim at supporting innovation and exchange of good practices. 'The Erasmus+ Key Action 2 Strategic Partnerships' support the development, transfer and/or implementation of innovative practices and promotes cooperation, peer learning and exchanges of experience at European level.' (DG EAC, 2015a). Given that pupil mobility is based on different strategic partnerships between schools, the aim of specific pupil mobility depends completely on the aim of its overall Strategic Partnership. Typically, Strategic Partnerships aim to improve inclusion and basic skills of pupils, so that pupil mobility is employed as a means for achieving these aims. Some Strategic Partnerships might also aim at decreasing early school leaving rates and improve cultural awareness and understanding.

As such, in contrast to Key Action 1 mobility, there is no overall aim that can be attributed to all pupil mobility. Hence, any kind of counterfactual impact analysis needs to assume a common aim across all Strategic Partnerships initiating pupil mobility.

There are three types of exchanges that involve pupils for Strategic Partnerships in the field of school education. First, within 'blended mobility' physical and virtual mobility are combined. Second, short term exchanges of groups of pupils cover pupils' mobility at host partner schools abroad. Both, blended mobility and short term exchange have durations between 5 days and 2 months. Third, long term study mobility regards pupils aged 14 years and older and covers durations of 2 to 12 months. Participation in long term pupil exchanges is more difficult to achieve within Strategic Partnerships, so that pupil numbers are very low (DG EAC, 2017. pp. 124-129)

Pupils participating in mobility are selected by the school, which participates in a Strategic Partnership.

5.3.1 Existing research

To the best of the knowledge of the authors, the impact of this action has not been evaluated yet. One relevant caveat in studying the impact of pupils' mobility is its heterogeneity ranging from short-term group exchanges to individual long-term stays in

host schools abroad. Short-term experiences might have an impact, but this impact is probably small and as such only possible to detect by using population data or investing into large expensive survey data collection. It could therefore be argued, that impact evaluation should generally focus on individuals' long-term mobility. However, the number of pupils taking part in long term activities is relatively low.

A second difficulty is the lack of individual data on which to evaluate the programme.

Third, given that the aim of pupil mobility depends on the 'Strategic Partnership' within which it is organized, the choice of outcome variable for causal impact analysis is not straightforward.

The ongoing evaluation of Erasmus+ plans to carry out a counterfactual impact evaluation of pupils mobility based on survey data that will be collected by the contractor in the current academic year. The proposal aims to collect the individual data at two different points in time (before and after the mobility, for a group of participants and a group of non-mobile pupils) which allows using a Diff-Diff approach in which the before-after programme differences between mobile and non-mobile pupils' outcomes are computed. This is an interesting approach which however will need to address a number of limitations on data compilation.¹⁷

5.3.2 Data

Any kind of data collections would depend on an agreement of the outcome variable to measure. Given that different Strategic Partnerships pursue different aims this is not straightforward to do.

For future evaluations, three different kinds of data sets could be considered:

- i) Exploit national administrative pupil data sets including information on **students' educational achievement.**

As discussed above, students' basic achievement might - but not necessarily is - the aimed outcome of specific strategic partnerships. The following possibility of data access is however only of importance, if DG EAC would like to measure the impact of pupil exchange on achievement. In addition, educational achievement is unlikely to change with short term pupil exchanges. As a consequence, the following data collection should only be considered if the focus is on long term pupil exchanges (2 to 12 months) of pupils aged 14 years or older.

A number of countries conduct regular national exams for pupils at different ages and collect information on pupils' results and other characteristics in a data base covering the entire population. The UK is an example, but also Scandinavian countries and some other central European countries have similar administrative data.

¹⁷ The average length of the considered mobility is very short (10 days). The estimated average effects, if any, will be necessarily very small. As such, the sample size would need to be very big to show that the small effects found are not just by chance of drawing a specific sample. This is however the major limitation of the planned evaluation. The contractor proposes a so-called 'cluster design', collecting first information on 50 schools and then students within schools. Cluster sampling lessens the precision of the estimates (increases the standard error) especially if the schools differ from one another (so-called intracluster correlation) and the school sample is small. This is the case with just 50 schools planned to be sampled. Given the combination of both, short mobility and lack of precision due to small samples, it will be very difficult to find a significant effect. It is also important to note that any kind of analysis would also need to be based on random samples (see Section 3).

National tests are implemented in specific grades only, but most of the countries conduct testing after age 14, which is the age requirement for participating in long-term individual pupil mobility.

However, this population data lacks an identifier for the pupils who took part in pupil mobility. This identifier could be obtained from DG EAC data, by merging individual information of students who were mobile with the country's administrative data.

The creation of such a data set would be restricted to one or several countries where access to administrative data is achieved and merging with DG EAC data successful. For example, the UK's National Pupil Data set includes information on pupils' achievement before and after age 14 and would therefore be an obvious candidate for such an analysis.

This option would provide a window of opportunity to conduct high quality counterfactual analysis on population data in the field of pupil mobility with low costs.

The advantages are:

- The use of high quality population data makes it also possible to detect causal effects that are small;
- nationally standardised test scores are available for the population of pupils in the country;
- the results/scores of these tests are highly credible if they are compulsory (as opposed to tests that are not compulsory which the pupils could fill in with low motivation and effort);
- these outcomes are already collected for administrative reasons.

The disadvantages are:

- Only some countries conduct nationally standardised tests;
- access to national data sets needs to be negotiated with countries' data holders (for the UK data will be available);
- merging of highly sensitive national pupil data with DG EAC data needs to be negotiated with the data holders;
- only a limited amount of outcome variables can be assessed, like educational achievement.

- ii) Use national data on educational attainment to measure early leaving rates looking at school level data. Second, one could focus on **educational attainment**, e.g. looking at drop-out rates, grade completion, and choices in the educational track of the pupils. The use of such a data set would be suitable, if DG EAC decided to focus on early school leaving rates as outcome variable for evaluating pupil mobility.

Such an analysis could be conducted at the individual level, comparing i.e. school leaving rates among pupils taking part and not taking part in exchange programs. However, if the focus were on individuals, it could be difficult to receive information on pupils' background (i.e. parental education).

A better approach would be to conduct the analysis at the school level thereby comparing schools' early school leaver's figures between schools who participated and did not participate in strategic partnerships using pupil exchanges.

Advantages:

- Early school leaving rate is an objective measure.

- This measure can be obtained from the administrative records of schools for the entire population of pupils. Also school level data should be accessible.
- These records are already collected by the competent national bodies for administrative reasons.

Disadvantages:

- Data availability should be checked and data access negotiated with the competent national bodies in the Member States.
 - Merging of the individual data with DG EAC data needs to be negotiated.
 - Early school leaving rates might not be the outcome variable of primary interest for DG EAC.
- iii) Conducting a survey on school pupils. If DG EAC would like to determine a specific set of outcome variables – like **understanding and attitudes to European values** – it is most reasonable to conduct a survey on pupils. For the survey design it is important that first the survey is conducted using a random sample design to insure representativeness of the data and second to choose a sample size that is sufficient to identify small effects especially if short term exchanges of pupils are examined (given that those are most likely to render small effects only).

The list of schools with pupil mobility is held by DG EAC. Schools would need to be addressed and asked whether they were willing to participate in the survey. Randomly selected schools with pupil exchange would need to provide a list of all students in classes in which pupil exchange took place. Using this list, a random selection of students (participating and non-participating in exchange) would need to be drawn.

The sample size would need to be calculated upfront, taking into account an estimated 'effect' of student exchange of the outcome variable of interest and the sampling design (clustering requests higher, stratification lower sample sizes). A rich survey would need to be conducted, that collects in-depth information on students' background (like gender, age, class repetitions, parental education, parental occupation, educational achievement, attitudes to school etc.) as well as a good coverage of outcome variables to be evaluated (depending on DG EACs interest).

5.3.3 Outcome variables

As discussed above, given that pupil mobility is embedded in Strategic Partnerships the aim of mobility depends on the objective of each specific partnership.

Nevertheless, even though partnerships might aim at different outcomes by using the means of pupil mobility, pupil mobility could generally lead to a homogenous outcome like lower early school leaving, better language skills, more awareness of European values and better basic skills. The three possibilities for accessing data given above assume such a common benefit of pupil mobility.

5.3.4 Methodology

The causal impact of pupils' mobility on educational achievement and attainment could be assessed by means of a Diff-Diff and PSM. The idea would be to compare the change in some outcome variable (e.g. the educational achievement) of pupils participating in mobility before and after the exchange, with the change in the same outcome of pupils

that did not participate in mobility. The change in the outcome of the non-mobile pupils represents the normal evolution in the outcome. It would be examined whether achievement for participants followed another steeper trajectory.

Diff-Diff could be combined with PSM in order to enhance the comparability between the treated group and the control group. However PSM requires a rich data set on individuals' background variables that potentially could be associated with the outcome variable (see Section 2). For example, the UK national pupil database (NDP) includes in-depth information on socio-economic background and school attended. It can be merged with regional information (like deprivation index by region). All these variables are likely to be associated with achievement. As such, the data set offers good information for taking selection-bias into account.

5.3.5 Value added

Given that there is currently no literature available on the impact of pupil mobility on any kind of outcome variable as well as the considerable budget allocated to pupil mobility, counterfactual impact evaluation in this area would be of high value. The fact that the action targets pupils makes it very interesting for an impact evaluation: mobility at the early stage of the educational path may potentially contribute to lessening educational inequalities and thereby improving children's educational outcomes and future opportunities. As long as the action is not examined the potentially great impact remains unknown.

5.3.6 Time frame and risk assessment

In order to collect any of the three data sets discussed above, more examination needs to be conducted on the availability of data for different countries and access to lists of schools and pupils taking part in strategic partnerships from DG EAC. This would be an important step for measuring the impact of the post-2020 programme.

i) The impact of pupil mobility on national achievement scores

Regarding an examination of the causal impact of pupil exchange on achievement, suitable data for the UK is available which could be accessed within 3 months of time. The analysis would be of greatest interest, if it could be also conducted in other countries (which is likely to be the case). In addition, DG EAC would need to provide student identifiers for those UK students who participated in student mobility.

If population data on students can be merged with DG EAC pupil identifiers, the analysis is very promising. Standard errors do not need to be calculated with population data, so that even small effects found are significant.

ii) The impact of strategic partnerships including pupil mobility on early school leaver rates

In some countries, school level information on early school leaver rates, children's composition and average achievement are available. If DG EAC could provide a list of schools who participate in strategic partnerships using pupil mobility this information could be merged into the school level data. The availability of both data sources would need further investigation.

iii) Collection of pupil survey data

Data documentation by the current evaluators' survey on pupils would be helpful to assess the need for conducting an additional student survey.

5.4 Mobility of vocational trainees

17% of the total Erasmus+ 2014-2020 budget has been allocated to vocational education and training (VET) considering both learners and staff. In 2014, more than 100,000 learners participated in mobility programs in companies or in vocational institutes (DG EAC 2015c). The duration of trainees' activity can last between 2 weeks and 12 months. The main aims of the activity are improving participants' performance and language skills, enhancing employability and career prospects, increase self-empowerment and self-esteem as well the awareness of the European project and European values so as to better prepare the trainees to the labour market.

Eligible participants are apprentices or VET students resident in the country of the sending organization. Graduates of a VET school or company may participate (within one year of their graduation). Participants are involved in trainings abroad, in an enterprise/organization or in a VET school.

5.4.1 Existing research

Research on trainee mobility is limited to some quantitative studies. Ecotec (2007) and PPMI (2011) use questionnaires which ask the mobile learners to evaluate the impact of mobility on a variety of outcome variables like language and intercultural skills, willingness to vote at European election, European identity, and professional skills. Results of these studies show that trainees who participated in mobility programs report among others an improvement in personal and language skills, a considerable benefit in terms of vocational skills and higher employability.

To the best of the knowledge of the authors no counterfactual impact evaluation has yet been carried out to assess the impact of mobility in the VET field.

5.4.2 Data

The feasibility of a counterfactual impact evaluation is hampered by lack of data. Since 2014, data on participants is available from DG EAC's Mobility tool database. Data on a comparison group of VET non-mobile learners needs to be gathered from other sources.

In principle, the following two approaches for getting access to data suitable for counterfactual impact analysis are feasible:

i) Survey on VET students

One possibility is to conduct a representative survey on VET students (mobile and non-mobile) belonging to institutions that participate in VET mobility. A list of these institutions would need to be obtained from DG EAC. Random sampling methods should be applied considering clustering by VET institutions (thereby facilitating data collection). Sampled VET institutions would need to provide a population list of VET students including those participating in mobility programs. Random sampling of participating and non-participating students from these population lists could take stratification of trainees into account.¹⁸ While

¹⁸ Stratification increases the precision of the estimate if the variables used for stratification are associated with variable of interest. For example, if socio-economic background of students is associated with their labour market performance, sampling of

conducting such a survey requires considerable resources especially if a phone or face-to-face survey is chosen, an advantage would be the possibility to include measures on the entire set of objectives of the action.

ii) Use of national administrative data on labour market records merged with DG EAC identifiers of VET students

An important objective of VET students' mobility regards enhancing employability and career prospects. Measures of these objectives include employment status and income, information regularly covered in administrative registers on employees or social-security registers typically held by national ministries for labour. Administrative data is more and more used for evaluating labour market policies. For the evaluation of VET trainees' mobility, it could be equally employed.

Administrative data source held by national ministries of labour is unlikely to identify previous VET trainees. In addition, the VET institute the VET trainee studied at will not be known. The latter is of interest, since it would be beneficial to include information on the uptake of VET mobility in VET institutions so as to apply an instrumental variable approach. As such, administrative data on labour market performance would need to be merged with administrative data on VET participants from VET schools. In addition, if VET school data does not include information on who participated in the mobility scheme, a further merging with DG EAC data on VET mobile trainees would be needed. While for the scope of this feasibility report, possible data sets available could not be investigated in greater detail, it might be worth exploring which data sets exist and are suitable for a counterfactual impact analysis of the action.

5.4.3 Outcome variables

If a survey is conducted, the choice of outcome variables can mirror the numerous outcome variables of the action. If administrative data are used, measures of labour market success will be evaluated.

5.4.4 Methodology

The main identification problem in assessing the impact of VET mobility is again selection bias: that is, mobile VET students are likely to be different from non-mobile ones in terms of a number of characteristics (such as motivation, skills) that are unobservable to the researcher and that are also likely to affect the individual labour market outcomes (see Section 2). In this context, the selection bias could be tackled with PSM, differences in differences or an instrumental variable approach.

If survey is used that is administered after mobility, the analysis is most likely confined to PSM. The collection of survey data could aim at covering a large number of trainees' background characteristics to make the independence assumption most probable (see Section 2). If a survey would be conducted before and after mobility and the main outcomes would be related to measures not related to the labour market (like i.e. cultural awareness), a differences in differences approach could be utilised.

If administrative data can be exploited, an instrumental variable approach could be possible. This identification strategy has been already used in the higher education field

VET students should be stratified by socio-economic background, so that the socio-economic composition of VET students is not varying by sample probability but reflects the population proportion exactly.

(e.g. Di Pietro, 2015). A good instrument could be the number of students participating in mobility in each VET institution and for each discipline.¹⁹

5.4.5 Value added

To the knowledge of the authors, no evidence on the causal impact of VET trainees' mobility exists. As such, it is not known whether the action works in achieving its aims.

5.4.6 Time frame

As discussed above, an in-depth study would need to investigate the availability of the following data sets: administrative data on labour market characteristics, DG EAC data on VET institutes who participate in mobility schemes and administrative data sets on trainees by VET institutes covered in the DG EAC list. For each country, the availability of administrative data sets on labour market characteristics and VET institutes will differ. As a consequence, we would estimate that it could take 2 to 3 months to investigate data availability.

In case a survey would be conducted, DG EAC data on all VET institutes participating in trainee mobility needs to be exploited. Countries legislation on sharing information on trainees of VET institutes might differ which would impact on the possibilities to draw random stratified samples. Also in this case in-depth information would need to be gathered about access to sampling frames or population lists of VET institutions and their trainees in different national contexts.

5.5 Master Loans

The Erasmus+ Master Loans grants EU-guaranteed loans to students who are willing to complete a full study programme at the Master level in another Programme Country. Applicants must hold a Bachelor degree (or equivalent first cycle higher education degree) and must have been accepted for a second cycle programme (Master or equivalent) at a recognised higher education institution abroad.²⁰ Note that no formal condition is set on the financial and economic conditions of the applicants.

In particular, the loans are provided to the students by banks participating in the action. In turn, participating banks are guaranteed by the European investment Fund (EIF) on behalf of DG EAC. The loans are up to 12,000 and 18,000 Euros for one-year and two-year Master programme, respectively. They offer favourable interest rates (lower than the market rate) and pay-back terms.²¹

The Erasmus+ Master Loans was launched in 2015. While writing this report, the action is operating in Spain, France and UK through participating banks. The action is however becoming available incrementally in more Programme countries. In terms of budget, 517 million have been allocated for the guarantees managed by the EIF. Although this represents only 3.5% of the Erasmus+ budget, it is expected to be boosted by the

¹⁹ The underlying idea is that the higher the mobility in a VET institute, the higher the probability of participation in mobility. At the same time, higher mobility in a VET institute should not affect the outcomes, controlling for VET institute and discipline of study.

²⁰ The Master must be completed in a Programme country that is located in a different country (i) from the country of residence of the applicants and (ii) from where he/she obtained the Bachelor degree.

²¹ E.g. a 'grace period' of up to one year before repayments begin and a further 'payment holiday' of up to one year.

funding from the participating banks during the current programming period, bringing the total budget available for loans to over 3 million Euros (2017 Erasmus+ Programme Guide, 2016, p. 117). In addition, it is estimated that 200,000 student loans will be supported in the period 2014-2020. This makes this action both interesting and suitable for an evaluation.

As part of the Key Action 1, the Erasmus+ Master Loans has the following objectives: improve learning performance and foreign language competence, enhance employability and career prospects, increase sense of initiative and entrepreneurship, self-empowerment and self-esteem, enhance intercultural awareness, promote active participation in society, awareness of EU values and motivation for taking part in future education or training. Therefore, natural candidates for outcomes variables are measures of subsequent labour market prospects of graduates, participation in civil/political society, and participation in education or training programmes. In addition, these loans are expected to be means to support equity and promote fairness for taking up post-graduate education within the EU (although they are not explicitly targeted to students with poor socio-economic background). Such speculation should be assessed empirically by means of a counterfactual impact evaluation by socio-economic background.

5.5.1 Existing research

A number of studies investigated the causal impact of financial support on student outcomes considering secondary education (e.g. Angrist et al. 2002; Angrist et al. 2006) and tertiary education (e.g. Nistico', 2014). Angrist et al. (2002) and Angrist et al. (2006) studied the impact of PACES, a programme implemented in Columbia that provided vouchers to low-income high-school students for enrolling in private schools. These vouchers provided a strong incentive for students to perform well at school since they were renewed conditional on grade promotion. The authors exploited the random assignment of the vouchers to evaluate the causal impact of the programme and showed that it improved educational achievement (measured by e.g. high-school graduation rate, grade repetition rate, test scores). Based on a similar setting, Angrist and Lavy (2009) evaluated the causal impact of awarding cash payments to low performing students in case of successful completion of an exam that is pre-requisite for enrolling in post-secondary education in Israel. This study found positive effect on the likelihood of college attendance for girls but not for boys.

In addition, Nistico' (2014) investigated the causal impact of PhD scholarships on the subsequent research productivity of PhD researchers based on Italian data. The author found evidence that receiving scholarships during the PhD increased the probability of entering a profession in research institutions and of publishing scientific articles.

Lastly, Chapman (2015) studied the impact of educational loans on the labour market outcomes of college graduates in the US. The author exploited the cut-off test score used to assign the loan according to the Merit Aid programme, so as to evaluate the causal effect of the programme on the early career outcomes of college graduates. The results suggested that college graduates whose tertiary education was financed through loans are more likely to enter in jobs that pay higher salaries than college graduates with scholarships (who are more likely to enter in lower paying jobs with presumably non-financial benefits). In addition, the study showed that college graduates with loans are less likely to have careers in education after four years, which is considered as proxy for starting careers in public service. This evidence, which refers to the US, has clear policy implications. For example, which policies may reinstate the incentives for college graduates with loans to pursue public service careers? These considerations may be relevant also for the EU and hence points at the need of assessing the impact of Erasmus Master Loans.

5.5.2 Data

Currently, no data is available for carrying out a counterfactual impact evaluation of Master Loans. Hence the evaluation of this action requires a preliminary phase of data collection.

In addition, since this action was launched very recently, an ex-post evaluation of the impact of this action cannot be carried out in the short-term for two main reasons: first, the sample size of participants is too small for a statistical analysis; second, the impact of the action will materialise at the completion of the Master programme at the earliest, and hence the appropriate outcomes of interest cannot be measured yet. However, it is wise to design the evaluation and the data collection in advance. Note that the set of countries to be covered in the evaluation will depend on the degree to which the programme will be expanded to other countries.

A counterfactual impact evaluation relies on a comparison between participants and non-participants that are as similar as possible, so that the outcomes of non-participants would serve as the counterfactual representation of what would have happened to the participants had they not taken the loans. Most reasonable would be to randomly select participants and non-participants from the list of successful and non-successful applicants from all participating banks which provide Erasmus+ Master loans. These individuals should answer to a survey, depending on resources, 6 months and/or 3 years after the completion of their post-graduate degree abroad. Questions could comprise all areas of interest for DG EAC, like attitudes, cultural values, subsequent educational career, and labour market success. It could also be considered, to conduct a survey prior to the mobility for all participants and non-participants. For instance, a possibility could be to require all applicants to fill in the questionnaire at the moment of application. Note that the prerequisite for collecting the data is that the participating banks keep record of names and contact details of all applicants so that the latter could be re-contacted for subsequent surveys.

In addition, since the Master Loan could eventually allow students with poor socio-economic background to undertake post-graduate studies, it would be interesting to see if this action is successful in reducing inequality and enhancing fairness across EU students. To do this it is important to collect detailed information about the loan but also about the economic and financial condition of the students as well as of their families.

5.5.3 Outcome variables

Given the nature of the treatment (that is the completion of post-graduate studies abroad) the outcomes variables to assess the impact of the Master Loans could be the same as the ones proposed to evaluate the impact of higher education mobility (see section 4.1).

5.5.4 Methodology

It is important to notice that any research would compare those individuals who received the grant with those individuals who would have liked the grants but did not receive it. Obviously it is quite likely that successful applicants are different in some characteristics compared to non-successful applicants (and if it is only that they are better in writing convincing grant applications). Therefore, great care needs to be applied to take selection into groups into account. As such, any survey would need to obtain information on individuals' ability as well as socio-economic background.

The best method to apply would be RDD. This requires loan providers to select successful applicants by administering an evaluation score to all applicants. Those applicants who reach a specific score are then selected. The RDD exploits the idea that

around the threshold score individuals are similar in all characteristics except loan eligibility. There would be a rationale to suggest that loan providers use a standardised applicants' rating scheme for making this type of analysis possible.

If RDD cannot be applied because the loans are not granted on the basis of a score, then one could rely on a Diff-Diff approach. However, the latter requires at least two rounds of data collection, one at the moment of the application (before the treatment), and another one sometime after the completion of the Master abroad. If instead the data will be collected only once after the completion of the Master study abroad (for both successful and unsuccessful applicants) then a PSM strategy could be used to examine the causal impact of the Master Loans.

5.5.5 Value added

Since this action is new, it has not yet been evaluated. We acknowledge that DG EAC has outsourced the 1st annual Student Beneficiaries Survey 2015 of the Erasmus+ Master Loan Scheme, which will feed into preparations for Erasmus+ post 2020.

Nevertheless, this programme would be very suitable for a counterfactual impact evaluation, provided that appropriate data on successful and unsuccessful applicants will be collected. Incorporating these data requirements already in the implementation of the Master Loans programme would enable researchers to carry out a sound counterfactual impact evaluation in a few years from now.

5.5.6 Time frame

As already mentioned, in the current programming period (2014-2020), it has been estimated that 200,000 student loans will be supported (DG EAC, 2015a). Given the expected high number of participants, the evaluation of the Master Loan could be planned for the final evaluation of the Erasmus+ that should be delivered by 30 June 2022. However, this requires being able to access the list of all applicants (successful and unsuccessful applicants) recorded by the participating banks.

If this evaluation would be of interest for DG EAC, the evaluation exercise would require adapting the setting so that the planned data collection could be incorporated in the implementation phase of the action (e.g. establishing that the participating banks would provide necessary personal information of applicants for running the surveys). This would require an upfront effort from DG EAC. Then, the JRC would plan and design the surveys. As for the evaluation of the EVS (Chapter 5.1), the actual data collection process would be outsourced and the JRC would supervise the data collection. Then the JRC would carry out the analysis in-house. For this specific proposal no precise time frame can be drawn at this stage.

5.5.7 Risk assessment

As for the case of the EVS project, the risk associated with evaluation of the Erasmus+ Master Loans are mostly associated with the difficulties of ensuring high quality data collection through surveys. A prerequisite for collecting the data is to be able to access contact details of all applicants to Master Loans.

Data quality will be ensured by looking at the following issues:

- decide the country coverage by maximising the representativeness of the country-specific samples with respect to their population;
- minimise low-response rate;
- design the questionnaires so as to maximise high quality responses.

Once the data collection will be properly carried out, there are no risks associated to the analysis.

6 Erasmus+ actions difficult to evaluate with CIE

The measurement of causal impact of Erasmus+ actions regarding education staff (school teachers, higher education staff, VET trainers and teachers involved in adult learning) are difficult as outlined below.

Mobility of education staff, whether it regards higher education, school education, vocational education or adult learning staff, can comprise two activities: either teaching periods in a similar partner institute or training periods which support the professional development of teaching staff. While funding of teaching staff mobility varies (it is lowest for teachers of adults) but can be substantial²² it is generally exceeded by funding of student mobility, giving it less importance for being evaluated.

The objectives of mobility of education staff generally aim at improving competences and understanding of teaching practices in the area of education the participant is active in. The objectives however also relate to specific and topical tasks, for example school teachers' mobility can also aim to gain further knowledge of teaching refugees.

The duration of mobility for higher education, school, vocational training staff and adult teachers can be between 2 days and 2 months (Erasmus Programme Guide, 2017).

To the knowledge of the authors, there is no literature available on the causal impact of mobility of teaching staff whether it regards school, higher education, vocational or adult learning staff. The existing literature relies on subjective evaluations of teachers participating in mobility (Ecotec 2008, PPMI 2011). Data has been collected through surveys (before and after participation) which provides information on participants' perception of the impact of mobility on their personal and professional development, the work, the learning process of students, and on the education institution as a whole. These types of studies do not allow for conducting counterfactual impact evaluations due to lack of coverage of non-participants. In addition, subjective measures might reflect respondents' biases. Given the lack of causal impact evaluation, it is not known whether staff mobility works in terms of reaching its objectives.

However, causal impact evaluation of teacher mobility is difficult due to the following reasons:

i) Most obvious outcome variable: teacher's performance

The biggest challenge we perceive is to measure the objective of improved teacher quality which is subject of an ongoing debate in the literature. Pupils and students are unlikely to be good evaluators of their teachers. A strand of literature focuses therefore on teachers' value added on students' long-term outcomes (Chetty, et al. 2014a; Chetty et al. 2014b). This approach measures teacher's value-added by computing it as average test-score gains for his (her) students, adjusted for differences across classrooms in student characteristics such as prior scores. Clearly this approach requires the availability of a huge amount of data and is therefore very cost-intensive as long as the level of analysis is the individual teacher.

One possible approach would be to merge administrative data from schools or universities (if they included test scores for students) with administrative data on participation in teachers' mobility (which would need to derive from DG EAC). Administrative data from institutes would allow constructing a measure of the quality of the teachers, following the aforementioned value added approach. Linked data should allow assigning a measure of teaching quality to each teacher and information on

²² In 2014 around 43 million Euro were allocated to financing school staff mobility involving more than 21,000 participants.

teachers' mobility. However, it would be a considerable effort to investigate and access such a data set, which could not be produced for adult teachers' mobility.

A more practical approach would be to not focus on the individual but the institutional level. For example, pupil performance in schools which participate in staff training events could be compared with pupil performance of similar other schools. If performance data could be collected longitudinally covering also the period before teacher mobility was introduced in schools, a Diff-Diff approach could be applied.

If the causal impact analysis should aim at measuring another outcome variable (like attitudes to European values) a survey could be conducted on a representative sample of teachers based on a cluster sampling design. This would require the following steps: select a list of schools/universities/Vet institutions having implemented staff mobility, contact them, and then select randomly a number of teachers within each school. A number of measures should be put in place so as to ensure high response rate. Then an appropriate measure of the outcome variable of interest (including a variable on whether the teacher has participated in the mobility scheme) would need to be developed.

ii) Short duration of the measure

The duration of staff mobility is between 2 and 2 months. As such, the effect of the mobility could be relatively small. Small effects can only be shown with bigger sample sizes, requiring large scale surveys.

7 Erasmus+ actions impossible to measure with CIE

Within the Erasmus+ programme, a number of activities cannot be measured by means of counterfactual impact evaluation approaches. The following major reasons have been identified to prevent the use of counterfactual impact methods; these actions belong to Key Action 2 and 3 of Erasmus+.

A first problem is represented by the short duration of mobility/exchanges financed by an action. The reason why short-term mobility/exchanges are not suited to be evaluated with statistical analysis is because their expected impact may be too small to be detected in the analysis. This applies for instance to Youth Exchanges and short-term EVS/ESC or short-term mobility/exchanges financed through Strategic Partnerships.²³

A second reason that may prevent the use of counterfactual impact evaluation for some actions is the heterogeneity of the financed activities. In such case, it is very difficult to define an outcome of interest against which to assess the causal impact of the action. An example is parts of the Strategic Partnerships, for instance actions decided in the framework of strategic partnerships of schools (except for long-term pupil mobility which has been discussed in Section 5.2).²⁴ The heterogeneity of activities makes it difficult to define as well as to measure adequately the outcomes of interest. In a similar way, Knowledge Alliances and Capacity Building projects are not suitable for a counterfactual impact evaluation due to the heterogeneity of the embedded activities.²⁵ The same reasoning applies to IT Support Platforms (such as eTwinning, the European Platform for Adult Learning and the European Youth Portal) which offer a wide range of activities targeted to different types of individuals.²⁶ All the aforementioned actions belong to Key Action 2. Also Support to Policy Reform (namely Key action 3 of Erasmus+) is impossible to be assessed through a counterfactual impact evaluation because of the heterogeneity of the activities involved.²⁷ Finally, sport related actions are not suitable for an impact evaluation because the activities involved are very heterogeneous (they entail collaborative partnerships supporting non-profit sports events).

For these actions, we believe that other approaches such as descriptive statistics on participants and the events could provide a picture of the development of the activities supported by this programme.

²³ Strategic Partnerships can be in the field of education, training and youth. In terms of participation, strategic partnerships of schools involved the highest number of participants in 2014, which amounted to more than 105,000 individuals (DG EAC 2014a). These actions entail exchanges that can be short-term or long-term and can involve pupils or staff.

²⁴ Strategic Partnerships of schools is the type of partnership with the highest number of participants.

²⁵ This action may comprise mobility of students, researchers and staff of different kind to the extent that this supports the main objective of the alliance.

²⁶ These programmes offer virtual collaboration spaces, databases of opportunities, communities of practices and other online services for teachers, trainers and practitioners, young people, volunteers and youth workers.

²⁷ It finances activities, such as transnational and national meetings, that foster the dialogue between youth and policymakers with the goal of improving the quality of education, training systems and youth policies.

8 Conclusion

Erasmus+ is a major EU policy which has been subject of a number of studies. Nonetheless, as part of the wider Better Regulation approach the Commission aims to evaluate the **causal impact of policies** to enable judgement about whether policies work and to draw policy conclusions for future planning. Counterfactual impact evaluation (CIE) is a core part of that approach and the most rigorous method in academic economics. However, CIE approaches require access to more complex data sets than other more basic non-causal evaluation methods.

This report looked at what kind of evaluations could be conducted for measuring the impact of the post-2020 programme. It was based on an examination of existing data and considerations on data collection possibilities for judging on whether a CIE is feasible for objectives covered by Erasmus+ policies.

Results which are also summarised in Table 1 in the Executive Summary are as follows:

The JRC can carry out causal impact evaluations on higher education mobility and therefore contribute to the preparations for Erasmus+ post 2020 on how to best measure its impact. Regarding **higher education students' mobility** high quality administrative data merged with survey data is available for three countries: Italy, Germany and the UK. The data is representative for graduates in the three countries so that results can be generalised to the graduate population of these countries. Even though the analysis would be based on a sub-set of EU higher education systems, they would be likely to serve to inform wider policy considerations across the EU as a whole.

The focus on causal impact can be complementary and improves earlier work which does not meet the same level of sophistication. The major contribution of the counterfactual impact analysis is: (i) to provide causal impact evaluation thereby enabling a judgement on whether the policy reaches its objectives; (ii) the use of administrative and well-documented survey data representative for the population of interest; (iii) to compare the impact of student mobility across the three countries Germany, Italy and the UK; (iv) to investigate how the impact of mobility changes by socio-economic status thereby examining the inclusion aspects of Erasmus+; (v) to examine if and how the impact of mobility changed during the last decade in which student mobility numbers grew and labour market demand was volatile; (vi) investigate if and how the impact changes for students who studied outside Europe (see chapter 4.1).

The report further identifies the following four other Erasmus+ activities that can be evaluated: European Voluntary Service (or future European Solidarity Corps), pupil mobility, Vocational Education Trainee mobility, Master Loans and Jean Monnet Chairs. Evaluations of these programmes would be possible only in the long term.

While this report examined the availability of data provided by data holders not aiming to evaluate Erasmus+, counterfactual impact evaluation can be facilitated considerably if data needs are taken into account during policy creation and implementation (as is advised in the Better Regulation Agenda).

A good candidate for doing so immediately is the Erasmus+ action Master Loan programme that was introduced in 2015. This policy's implementation phase is still ongoing and currently expanded to different countries. If banks distributing loans would fulfil some basic requirements in terms of data collection and data holding, counterfactual impact evaluation would be facilitated greatly at low costs so that results could be in place for the final evaluation of Erasmus+.

References

- Angrist, J, Lavy, V. "The effects of high stakes high school achievement awards: Evidence from a randomized trial". *The American Economic Review* 99.4 (2009); 1384-1414.
- Angrist, J, Bettinger, E, Kremer, M. "Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia". *The American Economic Review* 96.3 (2006); 847-862.
- Angrist, J, Bettinger, E, Bloom, E, King, E, Kremer, M. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment". *The American Economic Review* 92(5) (2002); 1535–58.
- Angrist, J, Pischke, J "Mostly Harmless Econometrics: An Empiricist's Companion", Princeton University Press, 2009.
- Archambault, E, Campbell, D, Gingras, Y, Larivier, V. "Comparing Bibliometric Statistics Obtained from the Web of Science and Scopus". *Journal of the American Society for Information Science and Technology* 60 (2009); 1320–1326.
- Becker GS. *Human Capital*. New York: Columbia University Press and NBER, 1964.
- Bosquet, C, Combes, PP. "Do Large Departments Make Academics More Productive? Agglomeration and Peer Effects in Research". *CEPR Working Paper 9041* (2013).
- Chapman, S. "Student Loans and the Labor Market: Evidence from Merit Aid Programs". *Northwestern University Department of Economics* (2015).
- Chetty, R, Friedman, JN, Rockoff, JE. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates". *The American Economic Review* 104.9 (2014a); 2593-2632.
- Chetty, R, Friedman, JN, Rockoff, JE. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood". *The American Economic Review* 104.9 (2014b); 2633-2679.
- Corporation for National and Community Service: *Still Serving: Measuring the Eight-Year Impact of AmeriCorps on Alumni*. Washington DC: Office of Research and Policy Development; 2008.
- DG EAC (2017): *Erasmus+ Programme Guide, Version 1*. Retrieved from http://ec.europa.eu/programmes/erasmus-plus/sites/erasmusplus/files/files/resources/erasmus-plus-programme-guide_en.pdf
- DG EAC (2016a): *2017 Work Programme of Erasmus+, 2017 annual work programme for the implementation of 'Erasmus+': the Union Programme for Education, Training, Youth and Sport*.
- DG EAC (2016b): *EVS factsheets and impact 2016*. Retrieved from http://europa.eu/youth/sites/eyp_web_production/files/evs_factsheet_and_impacts_apr_2016.pdf.
- DG EAC (2016c): *The Erasmus Impact Study Regional Analysis. A Comparative Analysis of the Effects of Erasmus on the Personality, Skills and Career of students of European Regions and Selected Countries*. Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/library/study/2016/erasmus-impact_en.pdf accessed 18 January 2017
- DG EAC (2016d): *The Erasmus Impact Study Regional Analysis - Main results in a nutshell*.
- DG EAC (2015a): *Master Student Loans factsheet*.

- DG EAC (2015b): *Erasmus+ Annual Report 2014*. Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/library/statistics/erasmus-plus-annual-report_en.pdf
- DG EAC (2015c): *Erasmus+ Annual Report 2014. Annex I Statistics*. Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/library/statistics/ar-statistical-annex_en.pdf
- DG EAC (2014a): *The Erasmus Impact Study. Effects of mobility on the skills and employability of students and the internationalisation of higher education institutions*. Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/library/study/2014/erasmus-impact_en.pdf accessed 18 January 2017
- DG EAC (2014b): *Thinking About Europe. Jean Monnet Support to European Studies*. Retrieved from http://bookshop.europa.eu/en/thinking-about-europe-pbNC0413015/;pgid=GSPefJMEtXBSR0dT6jbGakZD0000a965F_sS;sid=raV3dC3043h3YH WLIQPp00_RqnZxp2CXs9c=?CatalogCategoryID=6R8KABstikAAAEjvJEY4e5L accessed 18 January 2017
- DG EAC (2012): *Survey on Prison Education and Training in Europe—Final Report*. Retrieved from http://www.epeamalta.org/uploads/3/0/6/4/3064611/survey_on_prison_education_and_training.pdf accessed 18 January 2017
- ECOTEC (2008): *The School Education of Children of Occupational Travellers in the European Union*.
- Elia, L, Santangelo, G, Schnepf S. *Synthesis report on the 'Pilot projects to carry out ESF related counterfactual impact evaluations*. Centre for Research on Impact Evaluation (CRIE) Econometrics and Applied Statistics, 2015.
- Engel, C. "The impact of Erasmus mobility on the professional career: Empirical results of international studies on temporary student and teaching staff mobility". *Revue belge de géographie* 4 (2010); 351-363.
- Fabian G, Hillmann, J, Trennt F, Briedis, K. *Hochschulabschlüsse nach Bologna. Werdegänge der Bachelor und Masterabsolvent(inn)en des Prüfungsjahrgangs 2013*. DZHW Forum Hochschule, 2016.
- Fujiwara, T, Kawachi, I. "Social Capital and Health: A Study of Adult Twins in the U.S.". *American Journal of Preventive Medicine*, Vol. 35.2 (2008); 139-144.
- Groves, RM, Fowler, FJ, Couper, MP, Lepkowski, JM, Singer, E, Tourangeau, R. *Survey methodology*. New Jersey (US); John Wiley & Sons Publications, 2009.
- Wu, H. *Social Impact of Volunteerism*. Points of Light Institute, 2011.
- Kawachi, I, Berkman, L. *Social Capital, Social Cohesion and Health*. Chapter in *Social Epidemiology* edited by Berkman, LF, Kawachi, I, Glymour, MM. Oxford University Press, 2000.
- Kawachi, I, Kennedy, BP, Wilkinson, RG. "Crime: social disorganization and relative deprivation". *Social Science and Medicine*, Vol. 48.6 (1999); 719-731.
- Mincer, J. *Schooling, Experience, and Earnings. Human Behavior & Social Institutions N. 2*. New York; National Bureau of Economic Research, 1974.
- Mincer, J. "Investment in human capital and personal income distribution". *The journal of Political Economy* (1958); 281-302.
- Mitchell, K. "Student mobility and European Identity: Erasmus Study as a civic experience?". *Journal of Contemporary European Research* 8.4 (2012).
- Mühleck, K, Grabher, A., Hauschildt, K., Litofcenko, J., Mishra, S., Ryska, R., Unger M. and Zelenka. M. *Testing the Feasibility of a European Graduate Study Final report of the*

EUROGRADUATE feasibility study, 2016. Retrieved from https://www.eurashe.eu/library/mission-phe/EUROGRADUATE_feasibility_report.pdf, accessed 18 January 2017

Nistico, R. "The Effect of PhD Funding on Post-degree Research Career and Publication Productivity". *CSEF WP N. 362* (2016).

OECD (2014): *PISA 2012 Results in Focus: What 15-year-olds know and what they can do with what they know*. OECD Publishing. Retrieved from <http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf>

Oosterbeek, H, Webbink, D. "Does studying abroad induce brain-drain?". *Economica*, 78 (2011); 347-366.

Otero, MS. "The socio-economic background of Erasmus students: A trend towards wider inclusion?". *International Review of Education* 54.2 (2008): 135-154.

Parey, M, Waldinger, F. "Studying abroad and the effect on international labour market mobility: Evidence from the introduction of ERASMUS". *The Economic Journal* 121.551 (2011); 194-222.

PPMI (2011): *Interim Evaluation of the Lifelong Learning Programme (2007-2013)*. Retrieved from http://ec.europa.eu/dgs/education_culture/more_info/evaluations/docs/education/llp2011_en.pdf

Rodrigues, M. *Does student mobility during higher education pay? Evidence from 16 European countries*. Luxemburg: European Commission. doi 10: 95642, 2013.

Sorrenti, G. "The Spanish or the German apartment? Study abroad related outcomes and its recognition by the labour market". (2015) (<http://www.aiel.it/cms/cms-files/submission/all20150606125647.pdf> accessed 18 January 2017).

Structure for Operational Support (SOS), (1999), "Volunteers' Perceptions of the Impact of the European Voluntary service in their life". Brussels

"The economics of volunteering. Hiding in plain sight" *The Economist*. (12th September 2014, <http://www.economist.com/blogs/freeexchange/2014/09/economics-volunteering>)

What Works Clearinghouse (2016). Key Criteria Used in WWC Reviews of Single-Case Design Research. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_scd_key_criteria_011017.pdf.

Zinovyeva, N, Bagues, M. "The Role of Connections in Academic Promotions". *American Economic Journal: Applied Economics* 7(2) (2015); 264-92.

List of abbreviations and definitions

CC-ME:	Competence Centre on Microeconomic Evaluation
CIA:	Conditional Independence Assumption
CIE:	Counterfactual Impact Evaluation
DG EAC:	Directorate General Education and Culture
JRC:	Directorate General joint Research Centre
DLHE:	Destination of Leavers from Higher Education Records
Diff-Diff:	Difference-in-differences
EACEA:	Education, Audiovisual and Culture Executive Agency
ECHE:	Erasmus Charter for Higher Education
EVS:	European Voluntary Service
ESC:	European Solidarity Corps
ETER:	European Tertiary Education Register
HEGESCO:	Higher Education as a Generator of Strategic Competences
HIS:	Higher Education Information System
IV:	Instrumental Variable
ISCED:	International Standard Classification of Education
ISTAT:	Italian Statistical Office
JM:	Jean Monnet
NPD:	National Pupil Database
PISA:	Programme of International Student Assessment
PSM:	Propensity Score Matching
RDD:	Regression Discontinuity Design
REFLEX:	Research into Employment and Professional Flexibility
SOC:	Standard Occupational Classification
STEM:	Science, Technology, Engineering and Maths
VET:	Vocational Education and Training

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: <http://europa.eu/contact>

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by electronic mail via: <http://europa.eu/contact>

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: <http://europa.eu>

EU publications

You can download or order free and priced EU publications from EU Bookshop at: <http://bookshop.europa.eu>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see <http://europa.eu/contact>).

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.



EU Science Hub
ec.europa.eu/jrc



@EU_ScienceHub



EU Science Hub - Joint Research Centre



Joint Research Centre



EU Science Hub



Publications Office

doi:10.2760/262344

ISBN 978-92-79-68821-8