

ECONOMIC EVALUATION OF TRANSPORT PROJECTS AND POLICIES: METHODOLOGY AND APPLICATIONS

Part 1: Methodology for the cost-benefit analysis of transport projects and policies

September 1, 2020

Authors:

Ginés de Rus	University of Las Palmas de Gran Canaria, FEDEA
Javier Campos	University of Las Palmas de Gran Canaria
Daniel Graham	Imperial College (London)
M. Pilar Socorro	University of Las Palmas de Gran Canaria
Jorge Valido	University of Las Palmas de Gran Canaria

External advisors:

Per-Olov Johansson	Stockholm School of Economics
---------------------------	-------------------------------

This technical report was commissioned to the authors by the Independent Authority of Fiscal Responsibility (AIReF) for the Spending Review on Transport Infrastructure. It is funded by the EU via the Structural Reform Support Programme (SRSS). This paper reflects only the authors' view and does not imply a policy position of the European Commission or the AIReF, neither are responsible for any use that may be made of the information it contains. This technical report has benefitted from long and fruitful discussions with the staff at AIReF, in particular, Santiago Fernández, Milagros Paniagua, Ángel Martínez and Fernando de la Cámara.

CONTENTS

1. INTRODUCTION	5
2. THE ECONOMIC EVALUATION OF TRANSPORT PROJECTS.....	10
2.1. How to measure social welfare changes due to transport projects.....	10
2.2. Transport demand and the measurement of direct benefits of transport projects	18
2.3. Beyond consumers and producers: generalizing the cost-benefit analysis model	22
3. DECISION CRITERIA AND TOOLS FOR EVALUATING TRANSPORT PROJECTS.....	25
3.1. Decisions rules in the economic evaluation of transport projects	25
3.2. Discounting benefits and costs.....	27
3.2.1. <i>Defining net present values</i>	28
3.2.2. <i>The relevance of the social discount rate</i>	30
3.3. The treatment of risk in cost-benefit analysis	34
3.3.1. <i>Risk sources in project assessment</i>	34
3.3.2. <i>Dealing with risk in project assessment</i>	35
3.3.3. <i>Decision criteria using risk analysis</i>	37
4. PRICING AND INVESTMENT DECISIONS IN TRANSPORT PROJECTS.....	40
4.1. Pricing and investment when there are no other transport alternatives	40
4.1.1. <i>Investment decisions when there is no economic cost of public funds</i>	40
4.1.2. <i>Investment decisions when there is an economic cost of public funds</i>	43
4.2. Pricing and investment when there are several transport alternatives	44
5. THE OPPORTUNITY COSTS OF PROJECTS: MARKET AND SHADOW PRICES.....	49
5.1. The cost of the resources involved in a project	49
5.2. Market and shadow price of inputs	50
5.3. The opportunity cost of labour.....	54
6. ECONOMIC VALUATION OF NON-MARKET GOODS.....	59

6.1. Methods for valuating external effects in transport projects.....	59
6.1.1. <i>Techniques based on transport-related markets</i>	60
6.1.2. <i>Techniques based on hypothetical markets</i>	61
6.2. Environmental externalities and transport projects.....	63
6.2.1. <i>Noise</i>	64
6.2.2. <i>Effects on the landscape</i>	66
6.2.3. <i>Air pollution</i>	69
6.2.4. <i>Climate change</i>	71
6.2.5. <i>Other environmental costs of transport</i>	73
6.3. Accident costs and the statistical value of human life	73
6.4. Congestion as an externality	76
6.5. The value of time savings	81
7. INDIRECT EFFECTS AND WIDER ECONOMIC BENEFITS OF TRANSPORT PROJECTS.....	83
7.1. Direct and indirect benefits in transport cost-benefit analysis.....	83
7.2. Source of Wider Economic Benefits.....	84
7.3. Incorporating Wider Economic Benefits in transport appraisal	85
7.4. The link between agglomeration, productivity and transport provision.....	86
7.5. Valuing Wider Economic Impacts of agglomeration in transport appraisal.....	86
7.5.1. <i>Measuring change in agglomeration</i>	87
7.5.2. <i>Estimating the elasticity of productivity with respect to agglomeration</i> .	87
7.5.3. <i>Elasticity values used to assess agglomeration benefits in the UK</i>	90
7.5.4. <i>Quantifying the agglomeration benefits arising from transport schemes</i>	92
8. CASUAL INFERENCE FOR EX-POST EVALUATION OF TRANSPORT INTERVENTIONS.....	95
8.1. Introduction	95
8.2. Causal inference and ex-post evaluation: terminology and key concepts.....	96
8.2.1. <i>Terminology and key concepts in causal inference</i>	96
8.3. The potential outcomes framework for causal ex-post evaluation.....	98
8.4. Causal statistical models for ex-post evaluation.....	101
8.4.1. <i>Outcome regression</i>	101

8.4.2. Propensity score models	101
8.4.3. Difference-in-differences (DID)	103
9. GOVERNANCE AND INSTITUTIONAL DESIGN	106
9.1. The provision of transport infrastructures	106
9.2. The planning of transport infrastructures	108
9.3. The evaluation of transport infrastructures	110
9.4. The regulation of private participation in the provision of transport infrastructures 111	
9.5. Price regulation	113
ANNEX A. REVIEW OF SOME OFFICIAL GUIDELINES FOR COST- BENEFIT ANALYSIS IN TRANSPORT.....	116
A.1. Project definition.....	122
A.2. Methods to identify benefits and costs.....	123
A.3. Measurement and aggregation of benefits and costs.....	126
A.4. Decision criteria and the interpretation of results.....	127
A.5. Other issues.....	129
ANNEX B. A THEORETICAL FOUNDATION OF THE COST-BENEFIT ANALYSIS RULES.....	131
REFERENCES.....	136

1. INTRODUCTION

Transport improvements, as a consequence of both investment projects and policies, can be typically contemplated as exogenous interventions in transport markets, which move the economy from one equilibrium to another, commonly through the reduction of the generalized price (monetary price, time, and other disutility components) borne by transport users. Even projects aimed to provide new capacity and, of course, transport policies, such as subsidies to passengers, can be analysed as a reduction in the generalized cost of transport.

Although there are different reasons that, in principle, could justify such interventions from the point of view of public policy (e.g. increasing accessibility, improving safety, decreasing congestion, scarcity, or reducing negative environmental externalities), the question is not the existence of social benefits thanks to the public intervention, but whether these potential benefits are large enough to compensate the opportunity cost of the resources diverted from other uses to yield those benefits. This is the challenge of the economic evaluation of projects and policies,¹ whose main objective is to assess the changes in the well-being of the individuals directly or indirectly affected by the impacts of the project in primary and secondary markets.

Direct effects are easy to identify and measure as time savings, increases in safety and reliability, cost savings and so on. However, the complexity of the task is higher when there are significant effects beyond the primary market, the so-called indirect effects, particularly when the rest of the economy is not perfectly competitive, and presents, as it happens to be the case, distortions as taxes, externalities, unemployment and market power. The spatial nature of transport introduces another potential benefit due to changes in proximity of workers and the possibility of increases in productivity through different mechanisms, such as industrial reorganization or changes in land use. These are the so-called wider economic benefits, and the risk in this case is to confuse relocation with growth. Relocation occurs when some benefits of the project come from deviation of the economic activity somewhere else, while growth occurs when the project adds value to the economy.

There are different tools available to practitioners to address this challenge. The main ones are cost-benefit analysis (CBA), multicriteria analysis (MCA) and computable general equilibrium models (CGE).² These *Guidelines* are based on CBA, the most

¹ It is a common misunderstanding to identify the economic evaluation of projects (comparing social benefits with social costs) with their financial evaluation, where the focus is only on revenues and producers' costs.

² MCA is mostly used in evaluating projects under conflicting criteria. CGE is gaining status in transport and is particularly recommended in the case of megaprojects where some of the requirements for the

common methodology in the main supranational and national economic evaluation manuals,³ which can be defined as the quantification in monetary terms of the incremental changes in welfare derived from the implementation of a project with respect to a counterfactual (the economy without the project), with the ultimate aim of examining whether the society is better off after the intervention.

This procedure is not only useful *ex-ante*, but also when the project has been completed (*ex-post*), or when it has been running for a sufficiently long period of time. In these latter cases, the assessment is not about deciding whether to approve or reject the project, but whether it should be amended (provided that new information is available) or about drawing lessons that could improve future projects. In all three cases the focus of the evaluation remains on achieving the maximum welfare of the individuals. In this sense, the main objective of CBA becomes assisting decision-makers in prioritizing socially relevant projects according to their computable contribution (in monetary terms) to social welfare.

For both *ex-ante* and *ex-post* evaluation, and also for ongoing projects, the usual approach is to re-run the *ex-ante* CBA calculations feeding the model with observed rather than predicted values. Although this is quite informative about the actual economic effects of the projects and how these effects compare with the *ex-ante* analysis, the weakness of the approach is that the evaluation is performed under the same model, assumptions and simplifications, of the *ex-ante* evaluation.

Additionally, we consider a different approach to tackle the *ex-post* evaluation, based on causal inference. It consists of applying statistical models to data observed before and after transport interventions trying to estimate the incremental economic effect of the project. The approach is empirical and, although it avoids the aforementioned problems of CBA, it has to overcome other statistical requirements to uncover causal effects. Both approaches are linked by the same concern, to contribute to the understanding of government intervention in the economy through transport investments and policies. Hence, in these *Guidelines*, although we mainly follow the conventional CBA approach, we do not apply rules of thumbs from different sources. Our narrative on how the transport sector works, and how the government intervention affects welfare, will be supported by an analytical approach from which our rules and

application of CBA are not satisfied. However, the use of the CGE models, for the economic appraisal of projects, requires distinguishing between the incremental effect in welfare (the CBA approach) and the economic impact analysis of a typical CGE model. The measurement of effects on gross value added, or employment, as the main outputs of these models, has to be adapted to produce something that can be interpreted as a monetary measure of the change in welfare due to the project, as it is the case in CBA.

³ A review of some of these manuals is provided in **Annex A** of this document.

criteria of measurement are derived, always explaining the assumptions and conditions under which our results are valid.

In this context, and within the main task of providing the framework for a general analysis of the transport infrastructure investments in Spain, using well-established principles and accepted practices of CBA, the objective of these *Guidelines* is to establish the methodologies for a more detailed evaluation of two selected spending areas by producing three related parts. Firstly, **PART I** will include a general CBA methodology for the economic evaluation of any government intervention in transport, including an up-to-date review of the best practices, and adding an empirically based method of causal inference. In **PART II** this methodology will be specifically adapted for its application to the evaluation of railways investments, distinguishing between high-speed, and suburban rail projects (*cercanías*), in accordance to their particular features and the economic impacts.

Thus, the remaining of this report is devoted to **PART I**, concerned with the practical application of economic principles to project evaluation. The report departs from an analytical model which is presented in detail in **Section 2**,⁴ considering that transport projects, such as building new HSR or commuting lines, or transport policies, such as awarding subsidies to island residents to increase their mobility, can be interpreted as perturbations in the economy affecting the welfare of different individuals at different moments in time compared with the situation without the project, which does not necessarily mean the status quo but what would have happened in the absence of the project.

It is necessary to compare the world with and without the project, to recreate an alternative world, or the so-called counterfactual. CBA practitioners have then to solve two main problems. Firstly, they have to build the counterfactual, and this means to replicate the world without the project, a dynamic world that evolves without the perturbation introduced by the project. It is evident that this is not an easy task because the time period for this exercise may be quite long, and the values of key variables will change during the time length of the project, only some of them in predictable ways. Secondly, the practitioner has to imagine the world with the project, forecasting the main changes with respect to the counterfactual that he has previously created. The decision criteria and the procedures to make these decisions are discussed in **Section**

⁴ The model is completed in **Annex B**, where the reader can check where the final formulas to measure benefits and costs come from. Although we have tried to avoid being too formal, this annex provides a unified structure in which a basic set of ideas and principles of public economics allows the construction of a common thread unifying the different parts of the methodology and giving a sound support to the more simplified, and intuitive, graphical and verbal exposition in the remaining sections of the document.

3, whereas **Section 4** takes into account how pricing mechanisms may affect these investment decisions.

The expected impacts when the project is implemented are then the result of the comparison with the counterfactual: the more accurate is the construction of the counterfactual, the better is the estimation of the economic impacts of the project. Hence, it is important to present all the assumptions and the data used to complete this exercise. Transparency and *ex-post* evaluation can help to avoid both innocent errors and strategic misrepresentation. Taking this into consideration, **Section 5 and 6** focus on how to value direct effects, distinguishing between market and non-market goods, whereas indirect effects and wider economic benefits are defined and measured in **Section 7**.

Both as an alternative and complementary methodology to standard CBA approaches, **Section 8** includes a discussion on the causal inference approach, that aims to provide robust and verifiable empirical evidence about the economic effects of the government intervention. Finally, **Section 9** addresses the main challenges and incentive problems associated with governance and institutional design in the overall process of economic evaluation of transport projects.

PART II (in a separate document), contains the CBA methodology for the evaluation of railway projects, whose overall social value may be large in congested corridors (urban or intercity) with enough volume of traffic to compensate the fixed and external costs of this irreversible investment, but much lower in corridors with weak demand. This report follows the document *A general methodology for cost-benefit analysis in transport*, described in detail in **PART I**. We define the ex-ante CBA methodology for the evaluation of railway projects, particularly focusing on the different roles played by the different stakeholders in the industry. The report contains the relevance of demand projections in the adequate definition of rail project, and the description of the most relevant technical features of these projects, particularly distinguishing between HSR and commuting or suburban railways. The report includes the application of the CBA methodology to each of these rail projects types, providing a hypothetical example where the first-year demand required for financial and economic profitability is calculated. Finally, it includes a review of some of the most recent CBA rail guidelines and enumerates a set of variables and data sources needed for performing the CBA of rail projects.

PART III (in a separate document) is based on the idea that government intervention may be necessary to ensure an adequate level of connectivity and mobility of residents in non-peninsular territories. There are different policies that the government may use in order to achieve this objective, each of them producing different effects on the market. Thus, it is necessary to develop a CBA methodology to assess the consequences that each alternative policy may induce in the market, identifying

winners and losers, and carefully analyzing the key parameters that would allow the public policy to produce the desired effects in the market.

To do so, in **PART III** we develop an economic model to identify the key variables and economic agents that should be considered when evaluating different policies aimed for residents in non-peninsular territories. In the analysis, two extreme situations regarding the market structure are considered: either a situation in which there are so many airlines operating in the route that none of them has any market power (the perfect competition case), or a situation in which just one airline operates in the route and, thus, such an airline has all the market power (the monopoly case). Any other real situation that might be considered regarding the market structure is between these two extreme cases.

We prove that the effectiveness of any policy aiming at increasing residents' air connectivity strongly depends on the particular characteristics of the route, such as the level of competition, the proportion of resident passengers, the shape of residents' and non-residents' demand functions, airlines' operating costs, etc. Thus, any CBA aimed at evaluating the effects of transfers for residents has to be performed route by route, taking into account the particular characteristics of the route and the period of time. Empirical models that use aggregate data are not informative enough to distinguish the routes where the policy is being effective from those routes where the policy is producing important non-desirable effects in the market.

2. THE ECONOMIC EVALUATION OF TRANSPORT PROJECTS

2.1. How to measure social welfare changes due to transport projects

The CBA of transport projects is commonly applied using a few measurement rules that are simply derived from transport demand.⁵ In these *Guidelines*, as explained in the previous section, we follow the conventional CBA methodology, but the analysis of how government interventions affect welfare through its effects on the economy is supported by an analytical approach to infer the foundations of these measurement rules, explain their assumptions, and clarify the conditions under which they hold. The ultimate objective of quantifying the incremental changes in welfare resulting from public intervention in transport markets is to assess the change in the well-being of the individuals living in the society, and this involves calculating, in monetary terms, the magnitude of the potential (*ex-ante* CBA) or actual (*ex-post* CBA) gains compared with the opportunity costs of the resources diverted from other uses for the sake of the project. Our departure point is a society composed by m individuals, each of which has a utility (well-being) function denoted by U_h representing his preferences over a set of consumption activities. Ignoring altruism and other complications within the treatment of preferences, we will consider that the higher the value of U_h , the better for individual h .

At an initial equilibrium, the aggregate well-being of these m individuals is denoted by a social welfare function,

$$W = W(U_1, \dots, U_m), \quad (2.1)$$

which is increasing in individuals' utility: that is, the society is better off when the individuals are better off. However, the weight that each individual's utility has on social welfare may be different, representing that the well-being of some individuals may be more important for the society than the well-being of others in terms of income/wealth, regional concerns, etc.

In this basic setup, a government intervention, through a transport project or policy, moves the economy from the existing equilibrium to another one. Once this exogenous impact is implemented and the initial equilibrium is modified, the challenge is to

⁵ This section draws on the approach by Johansson (1993) for the CBA of environmental changes, and de Rus and Johansson (2019) for the measurement of the effects of transport projects. Our model, in this section and in **Annex B**, includes the explicit consideration of time in the generalized prices of goods, and the corresponding budget constraint given the time endowment of the individual. This approach allows a more rigorous derivation of the CBA rules for the practical evaluation of transport projects.

estimate the corresponding change in social welfare. Two questions need to be addressed: first, how is each individual affected by the project?; and second, how does the change in individuals' well-being affect the social welfare? To answer them, we can formally represent the change in social welfare as a result of the transport project as:

$$dW = \sum_{h=1}^m \frac{\partial W}{\partial U_h} dU_h, \quad (2.2)$$

where the first term in each product of the right-hand side of the equation is the social weight attributed to each person in the society, and the second reflects the change in the utility of the individuals once the project is implemented. Under the simplifying assumption that welfare weights are the same for all individuals and hence equal to one, we can focus on the evaluation of transport effects just as changes in individuals' utility:

$$dW = \sum_{h=1}^m dU_h. \quad (2.3)$$

When income distribution is optimal or society has at its disposal means for unlimited and costless redistributions, monetary gains and losses can be summed across individuals, as it is done in expression (2.3). However, redistribution is not necessarily costless since, for example, it might affect incentives in a negative way. In this case, the actual income distribution may not be far from the constrained optimal one. This means that the actual situation represents a kind of constrained optimum and possibly we can just sum gains and losses across individuals. This is also sufficient if relative prices are left more or less unchanged.⁶

For simplicity, let us just consider a representative individual whose utility function depends on the consumption of all goods and services produced in the economy: $U(x_1, \dots, x_n)$, where x_j represents the amount of good or service j , with $j = 1, \dots, n$. This representative consumer chooses the optimal set of consumption activities that maximize his utility, given his budget constraint, which denotes all the combinations of goods and services that may be purchased given his income and market prices. We assume that the consumer is also the ultimate owner of all firms in the economy and,

⁶ See Johansson and Kriström (2016) for a detailed explanation of the aggregation problems that may arise and the practical approaches.

thus, obtains income not only from the payment of labour inputs (wages) but from firm's profits too. Thus, the consumer's budget constraint is formally given by:⁷

$$\sum_{j=1}^n p_j x_j \leq \Pi + wl, \quad (2.4)$$

where p_j represents the market price of good or service j , Π denotes consumer's profits income, w is individual's wage and l represents working time.

Everyday life activities are time-consuming, but in the case of transport, this input is particularly important and should be explicitly included in the analysis because individuals make their travel decisions both in terms of market prices and the opportunity cost of the time required to make each trip. How can we introduce this specific feature of transport decisions in the budget constraint given by expression (2.4)?

Let us denote by \bar{l} the time endowment available for the consumer (for example, 24 hours per day, or 365 days per year), and by t_j the time required to consume good or service j . Then, the working time is the difference between the total time available for the consumer and all the time required in his consumption activities:

$$l = \bar{l} - \sum_{j=1}^n t_j x_j.$$

Note that if a transport project reduces travel time, on the one hand, the consumer will have more time to work which in turn will lead to the production of additional goods. On the other hand, the project may imply a cost, measured in terms of the monetary value of the goods that the consumer has to give up in order to implement such a project.

Taking into account the relationship between the time endowment and working time, consumer's budget constraint can be rewritten as:

$$\sum_{j=1}^n g_j x_j \leq \Pi + w\bar{l}, \quad (2.5)$$

⁷ See **Annex B** for details about the derivation of all the intermediate expressions and other underlying assumptions in the model.

where $g_j = p_j + wt_j$ represents the generalized price of good or service j and $w\bar{l}$ is the value of the individual's time endowment.

For example, in the case of air transport, g includes the monetary price paid (e.g. the airline fare, airport charges, etc.) and the users' time cost (access and egress time, waiting time and/or in-vehicle time).

The opportunity cost of travel time is the wage rate (w) in our model. In practice, determining the value of time becomes an empirical question since for some individuals (those who are willing to work more, but unable to find employment) the wage rate could overestimate their true opportunity cost, whereas for others the wage rate underestimates the value of their non-working time, when others nonmonetary benefits are associated with the job. In practice, the value of travel time is usually denoted by vt_j (and not just wt_j as assumed for simplicity in our model).⁸

Finally, it is worth highlighting that the price and the value of travel time might not be the only relevant parameters affecting consumers' travel behaviour. When the overall conditions of transport services matter (in terms of comfort, reliability, safety, etc.), an additional element of disutility should be added to the generalized price, which then would become $g_j = p_j + vt_j + \theta$, where θ covers this additional element of disutility.

Assuming that the budget constraint is binding, and the consumer's maximization problem reduces to:

$$\begin{aligned} \max_{x_1, \dots, x_n} \quad & U(x_1, \dots, x_n) \\ \text{s.t.} \quad & \sum_{j=1}^n g_j x_j = \Pi + w\bar{l}. \end{aligned} \tag{2.6}$$

The solution of the above maximization program yields the market demands of all goods and services produced in the economy, given by $x_j^* = x_j(g, y^g)$, with $g = (g_1, \dots, g_n)$ representing the vector of all generalized prices, and $y^g = \Pi + w\bar{l}$

⁸ There are several reasons why the value of time may empirically differ from the wage rate. This is the case when both work and travel affect utility directly (and not only the budget constraint, as in our model), or working time is unaltered by travel time savings. In those situations, the value of time of each individual depends on the sort of travel they undertake, that is, the time at which the journey is made, the characteristics of the journey (congested, repetitive, or free-flow), the journey purpose (commuting or leisure), the journey length, the mode of transport, or the size of the time saving (see Mackie *et al.*, 2001, for further details).

denoting consumer's maximum income, which is composed of the sum of profits income and the value of consumer's time endowment.

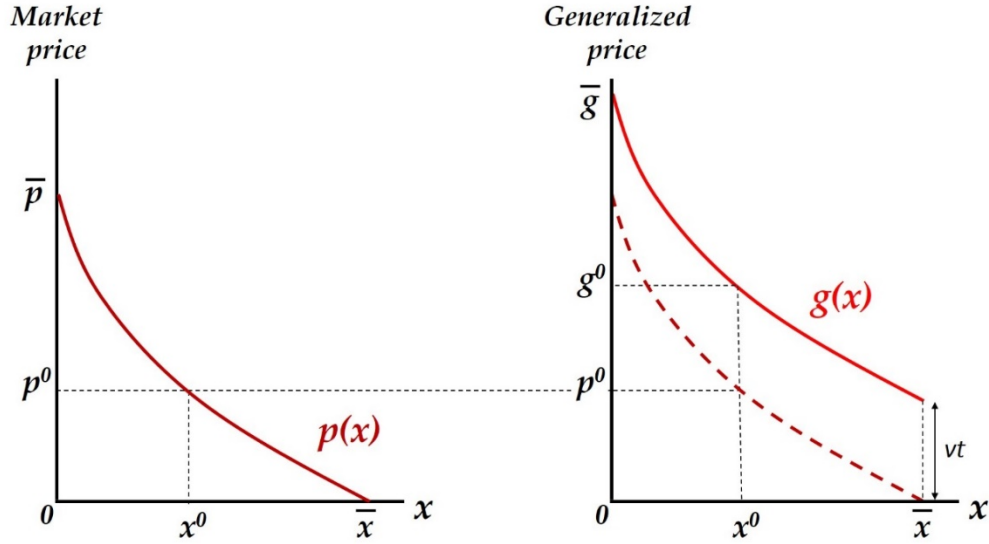
When the individual is maximizing his utility, the opportunity cost of one hour of time is the wage rate w , assumed as the value of time in our model because, since in the optimum, the individual is indifferent between consuming additional goods, or working more, and giving up the corresponding units of time. Hence, the hourly wage w , is the opportunity cost of time disregarding its final use (either leisure or consumption). This is the gate to the measurement of direct benefits of transport improvements: reducing the required time for transport, increases the time available for consumption in other goods or for working (though this may imply a cost, measured in terms of the monetary value of the other goods that the consumer gives up in order to implement the project).⁹

Figure 2.1 illustrates the relationship between the inverse demand function of a transport service in terms of the market price or the generalized price, assuming that the vertical distance between them, vt , is constant,¹⁰ and the disutility term θ is equal to zero. As explained above, in our model the demand function is the solution of the representative consumer's maximization program. In general, the market demand is the horizontal sum of the demand of all individuals' in the society. Thus, if x denotes the total number of trips demanded in the market, the (negative) relationship between the number of trips and the generalised price of transport services can also be interpreted as users' willingness to pay for them, in terms of the market prices and the value of travel time.

⁹ As later discussed in **Section 7**, in a more complex setup, where the spatial nature of transport activities is included in the model, the explicit treatment of changes in proximity and location could yield potential increases of productivity and the so-called 'wider economic benefits', and thus time savings (as measured in our model) would underestimate the social benefits of transport projects.

¹⁰ This implies that, for transport services, travel time is not affected by the number of users. Congestion will be dealt with in **Section 6**.

Figure 2.1. Transport demand function: market price vs. generalized price



These figures are commonly used to illustrate how to measure the direct benefits from transport investments. They come naturally from our model as described below.

By substituting markets demands, $x_j^* = x_j(g, y^g)$, in the (direct) utility function, $U(x_1, \dots, x_n)$, we obtain the consumer's indirect utility function:

$$U(x_1^*, \dots, x_n^*) = V(g, y^g), \quad (2.7)$$

which gives the consumer's maximal attainable utility when faced with a vector g of generalized prices and consumer's maximum income y^g . This utility function is called *indirect* because consumers usually think about their preferences in terms of what they consume rather than in terms of prices and income.

Let us now analyse the effects of a transport project, defined as an exogenous intervention that reduces the generalized price of transport and/or increases the number of trips, either via investment (e.g., an increase in capacity) or other policies (such as more efficient pricing, better management practices, etc.). The change in social welfare due to a transport project that implies a marginal increase in the number of trips can be measured in monetary terms as:

$$\frac{dW}{V_y} = \sum_{j=1}^n (g_j - wt_j) dx_j = \sum_{j=1}^n p_j dx_j, \quad (2.8)$$

where V_y represents consumer's marginal utility of income. According to this expression, the change in social welfare due to a transport project that implies a marginal change in the number of trips is equal to the difference between consumer's willingness to pay for those additional trips minus the value of travel time, that is, the market price. Note that if the transport project has a cost, some dx_j may be negative, representing the monetary value of production and consumption of other goods that the consumer has to give up in order to implement the project.

This is exactly what **Figure 2.1** shows. Let us suppose that x^0 in the horizontal axis corresponds to the situation without the project. This number of trips is associated with a particular consumer surplus (the difference between the consumer's maximum willingness to pay for those of trips and what he really pays), which is represented by the area above p^0 in the figure on the left hand side (equal to the triangle above g^0 in the figure on the right hand side). The height from x^0 to the demand curve is the willingness to pay (left figure) for an additional trip, or the generalized willingness to pay (right figure) for an additional trip. Hence, the value of an additional trip, i.e. dx in expression (2.8), is the generalized price of one trip, for a given level of trips (x^0), net of time costs, and this is simply the market price, $p = g - vt$, with $v = w$ in our model.

Alternatively, consider that the transport project implies a marginal reduction in the generalized price of transport. In this case, the change in social welfare is given by:

$$\frac{dW}{V_y} = -\sum_{j=1}^n x_j dg_j + dy^g. \quad (2.9)$$

If the change in the generalized price is caused by a marginal change in the market price, holding the travel time unaltered, the change in social welfare is zero. The reason is that, if all product and labour markets clear, a change in the market price without any time saving is just a transfer between consumers and producers. Moreover, there are no other additional welfare effects to be considered in the rest of the economy.

On the contrary, if the change in the generalized price of transport is due to a marginal change in travel time, keeping the market price unaltered, the change in social welfare is given by:

$$\frac{dW}{V_y} = -\sum_{j=1}^n x_j w dt_j. \quad (2.10)$$

In other words, the increase in social welfare due to a marginal reduction in travel time is equal to the value of the time savings ($dt_j < 0$) multiplied by the number of trips benefiting from that improvement.

This can be illustrated using the right panel of **Figure 2.1**. Assume that the equilibrium without the project is given by (x^0, g^0) . This number of trips is associated with a particular consumer surplus represented by the area of the triangle above g^0 . The base of this triangle is x^0 , and hence, when travel time decreases by dt in expression (2.10), consumer surplus increases by an additional area equal to $w dt$ (i.e. the change in the height of the triangle), multiplied by x^0 (i.e. the base of the triangle).

Expressions (2.9) and (2.10) are derived by considering marginal changes with respect to the situation without the project. When the effect of a transport project is not marginal, the change in social welfare can be directly approached as the change in consumer's utility with the project with respect to the counterfactual. In our model, the change in social welfare is, thus, given by:

$$\Delta W = \Delta V = V(g^1, y^{g^1}) - V(g^0, y^{g^0}), \quad (2.11)$$

where the superscript 1 indicates the values with the project and the superscript 0 denotes the values without the project and, thus, the social benefit of the project is expressed as the difference in the individual's utility with and without the project.

Although this utility is not measurable, expression (2.11) is very useful for a practical derivation of measurement. If the individual is asked how much money he is willing to pay to enjoy the benefits derived from the reduction in the generalized price of transport due to the project, we obtain a monetary measure of the change in utility. This is the so-called 'compensating variation' (CV), which can also be interpreted as how much money the individual would be willing to pay to have the project approved by the government. When CV is taken from the individual's income, he is indifferent between the situation with and without the project, as expressed by:

$$V(g^1, y^{g^1} - CV) = V(g^0, y^{g^0}). \quad (2.12)$$

If the project implies costs, the compensating variation does not only account for the benefits of the project, but also for the negative effects on utility derived from the diversion of goods and labour from other uses (the cost of the project). Therefore, the compensating variation represents the change in willingness to pay (WTP) due to the project benefits minus the willingness to accept for the goods and labour absorbed by the project. The net social value of the government intervention is then:

$$\Delta W = CV = \Delta WTP - \Delta \text{Resources}. \quad (2.13)$$

Interestingly, time savings, the main benefit in many transport projects, can be considered either as an increase in the willingness to pay or a positive change in resources. This is not important although, given the position of a generalized demand curve, the change in the generalized price of transport with the project increases the quantity, and so a change in willingness to pay of the induced demand quantity. For existing traffic, the willingness to pay (including time) has not changed and thus we can consider the value of time savings as a (positive) change in resources.

Suppose the representative consumer is asked for his willingness to pay for the transport project disregarding any effects on profits income. Then, the maximal willingness to pay, CV , as defined in expression (2.12), and the new partial one, denoted by CV^P are given by:

$$CV = CV^P + \Delta PS, \quad (2.14)$$

where ΔPS represents the change in firms' profits due to the transport project. If income effects are not significant, CV^P can be approximated through the change in consumer surplus (CS),¹¹ and then:

$$\Delta W = CV \approx \Delta CS + \Delta PS, \quad (2.15)$$

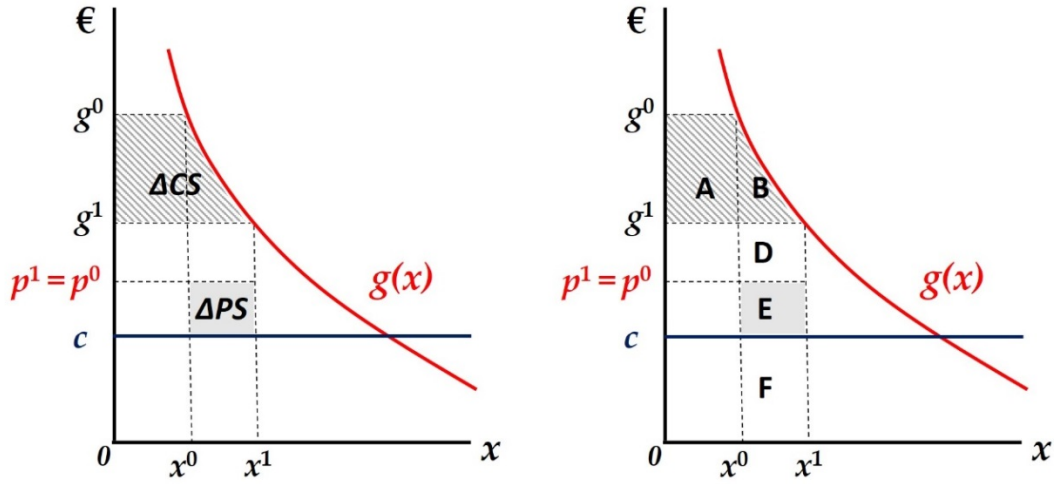
that is, social welfare changes can be approximated through the sum of the changes in the surpluses of the agents affected by the project.

2.2. Transport demand and the measurement of direct benefits of transport projects

After presenting the model that justifies the economic principles underlying the measurement of social welfare changes associated with transport projects, we now introduce the basic rules commonly used in CBA to measure these changes in practice. For the sake of simplicity, we will illustrate this analysis with a transport project without investment costs. In order to illustrate how to compute the change in social welfare due to such a transport project, let us consider the situation depicted in the left hand side of **Figure 2.2**, where $g(x)$ represents the market inverse demand function for a transport activity in terms of its generalized price. The initial equilibrium is (g^0, x^0) and marginal operating costs are constant and equal to c .

¹¹ The relative error of using the change in consumer surplus instead of CV^P is low if the elasticity of demand with respect to income, or the proportion of the change in consumer surplus with respect to income, is small enough (Willig, 1976).

Figure 2.2. Changes in social welfare due to a change in travel time



Due to the transport project, suppose that travel time decreases. Thus, the generalized price of transport is reduced to g^1 , whereas the market price remains at its initial level, $p^1 = p^0$; and the number of trips increase to x^1 . According to expression (2.15), the change in social welfare is the sum of the changes in consumer surplus and producer surplus, as represented by the corresponding areas, which can be easily calculated using the standard assumption of a linear approximation between the initial and the final generalized prices.¹²

Thus, the change in consumer surplus can be approximated by the area of rectangle $(g^0 - g^1)x^0$, which represents the benefits of the project for the existing users, and the area of triangle $(1/2)(g^0 - g^1)(x^1 - x^0)$, which represents the benefits for the new users (deviated and generated demand). Adding both areas, we obtain the so-called ‘rule of a half’:

$$\Delta CS \approx \frac{1}{2}(g^0 - g^1)(x^0 + x^1), \quad (2.16)$$

which, under the assumption of a linear approximation between the initial and the final generalized prices, is approximately equal to the stripped area in **Figure 2.2**.

¹² See Harberger (1967), Neuberger (1971) or Small (1999).

In this case, the change in producer surplus is simply the change in profits, or the difference between revenues and costs with and without the project:

$$\Delta PS = (p^1 x^1 - cx^1) - (p^0 x^0 - cx^0), \quad (2.17)$$

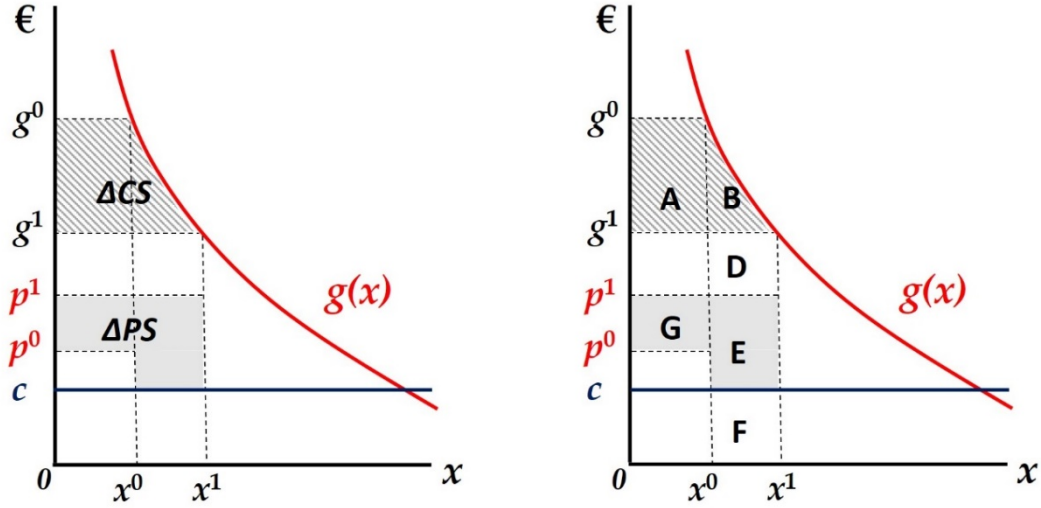
which coincides with the shaded area in **Figure 2.2**.

Alternatively, and using expression (2.13), the welfare effects of the project can also be measured through changes in willingness to pay and changes in the use of resources, as depicted on the right panel of **Figure 2.2**. The change in consumer's willingness to pay due to the project is given by the increase in the users' willingness to pay for the new trips, $(x^1 - x^0)$. By adding the reservation prices of the additional trips, we obtain **areas B, D, E** and **F**. From these areas, we have to subtract the resources required to obtain those benefits: **area D** (the value of the time spent in the new trips) and **area F** (the operating cost of the new trips). Finally, we must add a new benefit represented by **area A**: the value of the time savings for the existing trips. Thus, the change in social welfare due to the transport project is given by:

$$\begin{aligned} \Delta W &= \Delta WTP - \Delta \text{Resources} = \\ &= B + D + E + F - (D + F - A) = \\ &= A + B + E. \end{aligned} \quad (2.18a)$$

Notice that both approaches lead to the same result. Finally, it may be the case that the transport project is represented by the situation depicted in **Figure 2.3**. In this case, the project implies a reduction of the generalized price from g^0 to g^1 , but now this reduction in g is not just equal to the time savings, as the reduction in travel time is accompanied by an increase in the market price from p^0 to p^1 .

Figure 2.3. Changes in social welfare due to changes in time and market price



Applying the rule of a half, the change in consumer surplus in expression (2.16) is now represented by the striped area in the left panel of **Figure 2.3**. In this case, however, the change in market price increases revenues from existing trips, $(p^1 - p^0)x^0$ without affecting operating costs. The corresponding change in profits from the new trips is equal to $(p^1 - c)(x^1 - x^0)$, and the total change in producer surplus is given by the shaded area.

The right panel of the figure illustrates the measurement of the welfare changes according to expression (2.13). Again, the change in consumers' *WTP* due to the project is given by the additional willingness to pay of the additional trips, $(x^1 - x^0)$. Adding the reservation prices of all the new trips, we obtain **areas B, D, E and F**. From these areas, we have to subtract the resources required to obtain those benefits: **area D** (the value of the time spent in the new trips) and the **area F** (the operating cost of the new trips). Then, we have to add new benefits given by **areas A and G**. Both areas represent the value of the time savings for the existing trips. Thus, the change in social welfare due to this project is finally given by:

$$\begin{aligned}
 \Delta W &= \Delta WTP - \Delta \text{Resources} = \\
 &= B + D + E + F - (D + F - A - G) = \\
 &= A + G + B + E.
 \end{aligned}
 \tag{2.18b}$$

As expected, both approaches yield the same result again.

2.3. Beyond consumers and producers: generalizing the cost-benefit analysis model

Finally, expressions (2.11) to (2.15) can be generalized to include other stakeholders in the economic evaluation of transport projects. These agents are the workers, the taxpayers and the individuals affected by the externalities of transport. This latter group will be named 'rest of society'. It must be remembered that one individual can be in several groups and in fact it would be common to belong to all of them. The disaggregation is both for exposition convenience and later on for equity considerations.

Following Johansson (1993), the individual's indirect utility function is now given by $V(p, t, w, \Pi, \tau, z)$, where $p = (p_1, \dots, p_n)$ is the vector of market prices, $t = (t_1, \dots, t_n)$ is the vector of the time required for consuming each good or service, w is the workers' wage, Π are firms' profits, τ is a tax per unit of production, and z represents a set of natural resources.

In this setup, the change in social welfare due to a transport project (which implies a reduction in transport generalized price) is given by:

$$\Delta W = \Delta V = V(p^1, t^1, w^1, \Pi^1, \tau^1, z^1) - V(p^0, t^0, w^0, \Pi^0, \tau^0, z^0), \quad (2.19)$$

and using the concept of compensating variation, we have that:

$$V(p^1, t^1, w^1, \Pi^1, \tau^1, z^1 - CV) = V(p^0, t^0, w^0, \Pi^0, \tau^0, z^0), \quad (2.20)$$

with:

$$CV = CV^P + \Delta PS + \Delta WS + \Delta GS + \Delta RS, \quad (2.21)$$

where CV^P can be approximated by changes in consumers' surplus, ΔPS is the change in producers' surplus, ΔWS refers to the change in workers' surplus, ΔGS is the change in taxpayers' surplus, and ΔRS is the change in the surplus of the 'rest of society'.

The change in consumer surplus is calculated with the rule of a half using market prices (including taxes) as in expression (2.16). However, the change in producer surplus has to be modified to include taxes.

$$\Delta PS = (p^1 x^1 - c^1 x^1 - \tau^1 x^1) - (p^0 x^0 - c^0 x^0 - \tau^0 x^0). \quad (2.22)$$

The case of workers' surplus could be important in the evaluation of policies such as the privatization of public companies. In general, and assuming that the project under evaluation does not change wages (or they only change marginally), the change in

workers' surplus will come from the new workers required by the project. For simplicity and leaving the treatment of labour costs for **Section 5**, the surplus of a particular worker is equal to the wage minus his opportunity cost. It is worth recalling here that the opportunity cost of the worker does not necessarily coincide with the social opportunity cost of labour. The strict application of one of the two approaches contained in (2.23) reduces the risk of making mistakes.

The surplus of the rest of society is the valuation of the externalities minus the compensation received (e.g. payments of the airport authority to soundproof the windows), as later discussed in **Section 6**.

Finally, adding all the surpluses, the income transfers cancel each other and the result is again equal to the change in willingness to pay minus the value of the diverted goods and labour from other uses and the negative (or positive) external effects:

$$\begin{aligned}\Delta W &= \Delta CS + \Delta PS + \Delta WS + \Delta GS + \Delta RS = \\ &= \Delta WTP - \Delta \text{Resources}.\end{aligned}\tag{2.23}$$

From expressions (2.19) to (2.23) the individual provides a value (CV) which measures in monetary terms the change in utility once the project is implemented. When the project is financed with distortionary taxes (e.g. income or indirect taxes) the lost surplus in the economy has to be considered in the evaluation of the project. The tax increases the price and reduces the quantity and, though tax revenues collected by the government are a mere transfer, the net value of the lost output is an extra cost of the project. This is the so-called deadweight loss or excess burden of the tax and, in principle, should be included in the answer of our representative well-informed consumer revealing his willingness to pay for the project.¹³ Nevertheless, when the practitioner uses expression (2.21) what he is doing is to measure time savings, operating cost savings, increases in safety and so on, without asking the individual anything like the CV in expression (2.20). Hence, the marginal cost of public funds should be included in the evaluation of those transport projects that do not generate enough revenue to cover their costs.

The distinction between different agents in expression (2.23) does not mean that they are the final beneficiaries of the transport improvement. The existence of fixed factors, such as land, though it does not change the value of the final result of equation (2.23), may completely modify the distribution of the social surplus. It is well known that land can capitalize most of the benefits of transport improvements. In the case of an

¹³ The deadweight loss depends on what tax is raised or what activity has been crowded out to finance the project (see Johansson and Kriström, 2019).

infinitely elastic supply of homogeneous workers, the surplus of each group in expression (2.23) would be zero and the landowners would take the total surplus through higher land prices. This leads to a practical conclusion: it is easier to calculate the change in willingness to pay and the change in resources than the apparent distribution of the net surplus.

Furthermore, Collier and Venables (2018) have shown that with heterogeneity, both in labour productivity and demand for housing, workers can get a significant part of the surplus. The implication for the economic evaluation of transport improvements is that although the project increases the land values around the locations affected by the improvement, only in some extreme cases the increase in land values would reflect the total benefits of the projects because a share of those benefits are captured by workers.

Thus, the conclusion that transport benefits could be measured in a competitive land market when this market is not affected by bubbles or any other exogenous factors, only holds under some restrictive conditions. What is true, in any case, is that the practitioner should be very careful avoiding the combined use of the three possible approaches: change in surpluses, change in willingness to pay and resources, or the increase in land prices.

The procedures for the measurement of the welfare effects of transport improvements obtained in this section ignore the possibility of significant indirect effects, beyond transfers and relocation, and even the presence of wider economic benefits. The spatial nature of transport introduces other benefits from increases in productivity through different mechanisms, such as industrial reorganization or changes in land use. In any case, the qualifications about the calculation of the social surplus using expression (2.23) still apply. Both indirect and wider economic benefits are discussed in **Section 7**.

3. DECISION CRITERIA AND TOOLS FOR EVALUATING TRANSPORT PROJECTS

3.1. Decisions rules in the economic evaluation of transport projects

CBA is the main tool that help economists make rational decisions about transport projects. These decisions are generally made by identifying, valuing, and aggregating benefits and costs of different nature, summarising them into a single value that comprises all the relevant information associated with a particular project under evaluation. The comparison process then requires two elements: determining the benchmark (which is related to project definition and whether the CBA is performed before, after or during the project), and normalising the elements being compared, which involves an intertemporal homogenization of benefits and costs.

In this context, making rational decisions simply becomes a search for the best feasible allocation of resources (i.e., Pareto efficiency), and the immediate corollary is that society should only favour Pareto improvement transport projects, in which social welfare is increased because at least one individual is better off without making anyone else worse off. However, there are winners and losers. Building a new road or improving an existing one, for example, reduces travel time for some people, but also affects the well-being of residents in adjacent areas and may disrupt overall traffic patterns, causing congestion.

A Pareto improvement in these cases requires that anybody who would otherwise lose due to the project is fully compensated for his losses. This rarely happens and the criterion is too demanding for practical use in the economic evaluation of projects. Even the best project could make some individuals worse off. The criterion followed in practice is the Kaldor-Hicks compensation criterion, that states that a project should be adopted if the winners could compensate the losers and still be better off. This criterion involves that total benefits outweigh total costs. It does not require actual compensation and for this reason is also called the potential compensation criterion.

This principle can be directly translated to the evaluation of transport projects by comparing, in present value terms, the flows of total benefits (B) with their total costs (C). The net social present value (NPV_S) is in fact the potential compensation criteria. A positive NPV_S is indicating exactly that the winners could fully compensate the losers and some net gains are still available for the winners. Given that the assessment is made from the point of view of society as a whole, the comparison includes all social benefits and social costs of all the individuals affected by the project, even if some of them happen to be in other secondary markets beyond the primary market.

The decision-maker may be not only interested in the results of the NPV_S but also in the financial viability of the project (the change in the producer surplus of the public sector). This indicator is the NPV_F , where only revenues and producers' costs are considered. The NPV_F should be presented with the NPV_S because it shows the degree of financial sustainability of the project. Sometimes the availability of public funds is limited and the more self-financed the project, the lower the taxpayers' contribution and the need of additional taxation.

One of the most salient features of the CBA of transport projects is that their social and private benefits and costs often differ, but both are relevant in the decision process. As discussed in **Section 2**, a main part of the transport users' gains when their generalised price decreases, as a result of improved travel conditions (due to faster vehicles or less congested infrastructures, for example) is the value of time savings. These users' benefits are partially included in producer surplus if transport firms are able to raise the price to capture part of the time savings. Similarly, the social opportunity cost of the inputs required for the project is not necessarily the market price of those inputs. In the case of labour, for example, if unemployment is high, the social opportunity cost of employing additional workers may be lower than their market price. Moreover, the project may affect non-marketed goods, changing the level of externalities such as pollution or noise.¹⁴

Since these benefits and costs are the result of changes in market conditions which are predicted to occur if certain actions are undertaken, transport projects must be always defined in incremental terms, i.e. comparing what happens in the situation with the project with an alternative without the project. A strict 'do-nothing' counterfactual assumes, for example, that in the absence of the project, no action takes place at all, which may be suitable for capacity rehabilitation projects. More frequent is, however, a 'do-minimum' scenario, which assumes that there will be sufficient investment or any other action to keep existing operating capacity in the future, whereas a 'do-something (else)' scenario would consist of any other alternative approach to meet the project objectives.

Once all the benefits and costs have been properly measured in incremental terms, applying the potential compensation criterion, the economic evaluation of transport

¹⁴ How benefits and costs are valued mainly depends on the existence or not of markets for them. This will be discussed in detail in **Section 5** and **Section 6**.

projects is straightforward, and the corresponding decision criterion can be summarized into two simple rules:¹⁵

1. Any project with a positive social net present value ($NPV_S > 0$) should be accepted.¹⁶ If this happens for different (but comparable) projects, they all should be accepted in descending order of their NPV_S . If the projects represent mutually exclusive, but comparable alternatives, only the one with the highest NPV_S should be accepted.
2. If the financial net present value is negative ($NPV_F < 0$) and the decision-maker faces budget constraints, the project should be: (a) revised, facing the trade-off of improving the NPV_F and reducing the NPV_S (through changes in prices or quality, for example), or (b) delayed, if better demand or costs conditions are expected in the near future, or otherwise (c) rejected. Additional and more complex decision schemes can of course be devised when risk is introduced.

In sum, the economic and financial approaches in the evaluation of projects address two different but related questions. The first one is whether a project should be carried out or not at all, considering its contribution to social welfare. The second refers to the project ability to generate revenues above its costs, which is usually an important element for the viability of the project. The financial side is also important for the type of private participation in the construction, maintenance and operation of the project. Finally, as noted above, the financial evaluation is also essential to ascertain the project implications for public finances and the its effects on other government policies.

3.2. Discounting benefits and costs

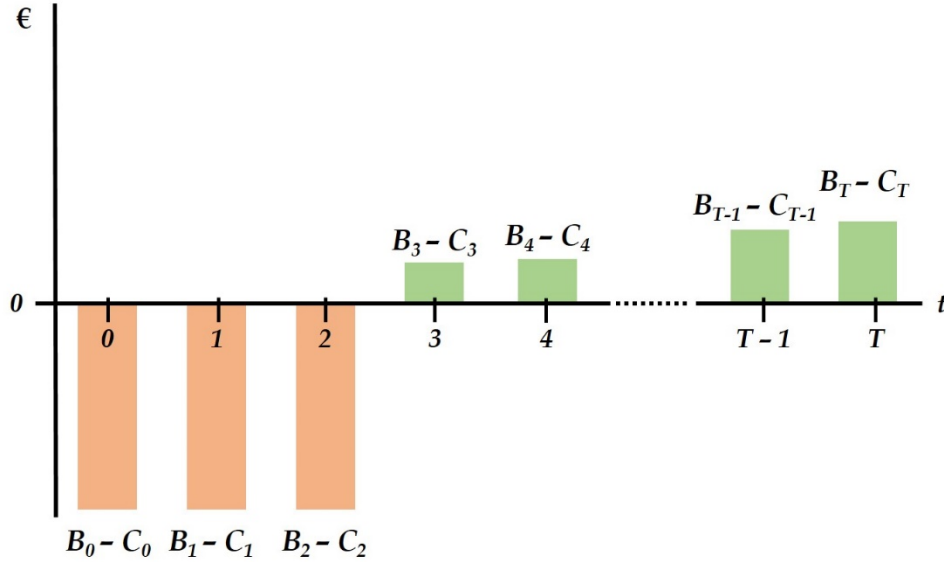
The effects of transport projects are generally distributed over discrete periods of time (e.g., years). They begin with an external intervention in a transport market, that may consist of an investment that covers several periods (e.g. greenfield infrastructure projects) or a one-off policy (e.g. changing the speed limit). Then, a stream of benefits and costs follows during the project life (T periods), which depends on the technical and economic characteristics of the project.¹⁷

¹⁵ Note that these rules are valid both for *ex ante* and *ex post* evaluations. The only difference is that in the first case both situations are hypothetical, whereas, in the second, the ‘with-the-project’ situation has already happened.

¹⁶ This decision is made regardless of the value of the NPV_F . However, if $NPV_S < 0$ the project must be always rejected because its social costs exceed its social benefits and, therefore, it does not pass the potential compensation criterion.

¹⁷ If the evaluation horizon is shorter than the project life, any residual value is added as a net benefit at T .

Figure 3.1. Time profile of a transport project: undiscounted values



As an example, **Figure 3.1**, shows the time profile of a project, where net values of benefits and costs, $B_t - C_t$, are negative at the beginning and become positive after the construction period, generating benefits which grow overtime.¹⁸ Since individuals do not assign the same value to a monetary unit regardless of when it is received, we cannot simply add these net values over time: they must be homogenized and compared from the same reference point. Although the choice of this reference is arbitrary, initial year ($t = 0$) is customarily preferred, and benefits and costs are then discounted and added into a *NPV*.

3.2.1. Defining net present values

As discussed above, when all social benefit and costs are included in the evaluation, the social net present value is equal to the sum of the change in social surplus or the sum of changes in willingness to pay and changes in resources. Hence, the *NPV_S* is the indicator of the economic profitability of the project under assessment:

$$NPV_S = \sum_{t=0}^T \delta^t (B_t - C_t).$$

¹⁸ In this example, investment costs are €150 million, equally distributed in years 0, 1 and 2. Net benefits start as €10 million in year 3, then growing at an annual rate of 2% during the rest of the project life

Conversely, when only the stream of revenues (px) and producers' costs (cx) are considered,¹⁹ from the point of view of producer surplus, as described in **Section 2**, the indicator of the financial profitability is obtained:

$$NPV_F = \sum_{t=0}^T \delta^t (p_t x_t - c_t x_t) .$$

In both cases, the discount factor (δ^t) represents the relative weight of each period. It generally adopts the exponential expression

$$\delta^t = \frac{1}{(1+r)^t} \leq 1 ,$$

where $r > 0$ is the social discount rate, which implicitly assumes that monetary units in the present yield a higher utility than future ones, and these are less valued the further they are. The divisor, $(1+r)$, can be interpreted as the opportunity cost of one unit of present consumption as consuming one unit today implies to give up 'one plus the discount rate' next year. This is a property of exponential discounting that could penalize projects whose benefits are realized in the long-term and favour projects with huge costs in the distant future.²⁰

CBA is concerned with changes in real values and inflation changes the value of monetary units over time. Consequently, the practitioner has to decide whether intertemporal comparisons of benefits and costs should be performed in nominal or real terms. The answer in fact is very simple: since the objective of CBA is simply to account for changes in welfare, the evolution of money values is irrelevant: what matters is the change in real benefits and costs. Nevertheless, although it is indifferent to apply CBA in transport projects using current values or constant ones (by deflating them to a base year), in some cases it may be preferable to use nominal terms, as this is how values are often reported or calculated. This approach also tends to be easier to interpret, particularly from a financial viewpoint and in projects with private participation. Whatever the selected approach, the evaluation must be consistent. If benefits and costs are expressed in nominal terms, a nominal discount rate (r_N) must be used. If they all are expressed in monetary values of the base year, the real discount rate (r_R) must be used. If φ represents the rate of inflation:

¹⁹ For simplicity, we assume constant marginal operating costs, as in **Section 2**.

²⁰ In these cases, a hyperbolic discount factor, $\delta = 1/(1+rt)$, with a slower decrease in the marginal rate of time preference, may be preferred, although it is not common in the transport sector. Some guidelines recommend a declining long-term social discount rate to give more weight to the benefits and costs of future generations (see, for example, HM Treasury, 2018).

$$r_R = \frac{r_N - \phi}{(1 + \phi)} .$$

The same reasoning applies when benefits and costs are valued in different currencies: they must be homogenised using the same exchange rate.

3.2.2. *The relevance of the social discount rate*

The social discount rate is a key parameter in the economic evaluation of transport projects. Its value does not only change the way benefits and costs are compared over time for the same project, but it can also change the ranking of different projects. **Figure 3.2** represents the stream of benefits and costs considered in the example of **Figure 3.1**, but now discounted at two different discount rates. Note that the higher the value of r , the lower the value of the discount factor, and hence the lower the value of the yearly net benefits. In the example represented in these figures, it can be guessed that the sum of the discounted benefits is large enough to compensate the construction costs when the discount rate is 3.5% ($NPV > 0$), but insufficient in the case of a 7% discount factor ($NPV < 0$). Therefore, a project with the same benefits and costs may be socially desirable depending on the social discount rate. This example illustrates the key role of the social rate of time preference in the selection of projects.

There may also be cases where changes in the discount rate modify the projects ranking, as showed in **Figure 3.3**, where the $NPV_A(r)$ function corresponds to a *Project A* and $NPV_B(r)$ corresponds to a *Project B*, with lower net benefits at its origin, but higher at the end. *Project A* is preferred only if $r < r^*$.

Figure 3.2. The effect of different discount rates on the same project

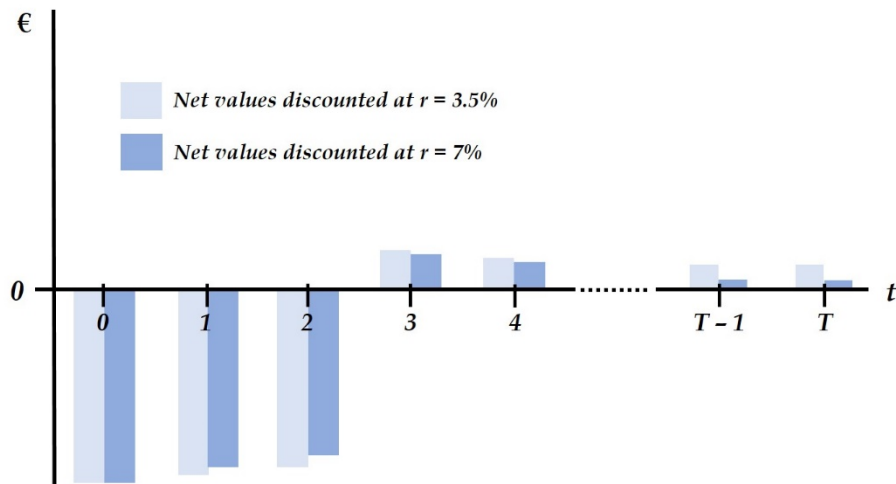
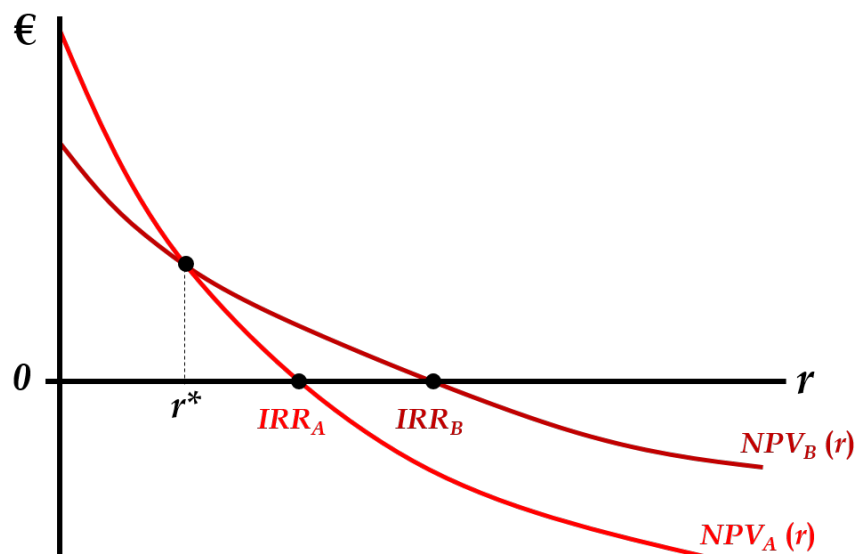


Figure 3.3. The effect of different discount rates on project ranking



The figure shows how the NPV decreases as r increases. Eventually, the NPV is equal to zero. At this point the value of r is called the internal rate of return (IRR). This value is another indicator of the social profitability of the project and allows to compare different projects with independence of their sizes. The rule to approve a project then becomes $IRR > r$, but despite this simplicity, the NPV remains as the preferred decision tool because the IRR has, at least, two technical problems: there may exist more than one solution to the $NPV(IRR) = 0$ equation, and it is not always the highest IRR project the one with the greatest NPV .

A final issue related to discounting that may be of interest for transport projects is the case when a decision has to be made between mutually exclusive projects that address the same problem, but with different lifespans. Comparability is a necessary condition for ranking these projects, and this may require a previous process of homogenization. Two procedures to make the projects comparable are commonly used. The first one is to consider a (fictitious) project of equivalent duration.; for example, if $T_A = 0.5T_B$, we could compare *Project B* with a ‘project’ consisting in performing *Project A* twice (the second time just replicating the first one just after T_A). The other procedure is to calculate the equivalent annual net benefits for each alternative, defined as

$$\hat{B} = \left[\frac{1}{S(T)} \right] NPV ,$$

where $S(T)$ is the discounted sum of a unit of net benefits during T years.²¹ For a given r , this expression allows to compare NPV s calculated for mutually exclusive projects of different duration intended to solve the same problem.

A positive NPV is a necessary condition for accepting a project but it is not sufficient. Even in the case of a single project with a $NPV > 0$ facing the ‘accept-reject’ decision, the practitioner should consider the optimal timing of the project. When the decision is not subject to a ‘now or never’ constraint, the investment is irreversible and there is uncertainty over the future rewards, the benefits and costs of postponing the project have to be considered. The circumstances that may justify postponing a project (when this is technically feasible) are varied. A first reason, for example, is that demand is growing and the social benefits derived from meeting demand in the first year do not outweigh the opportunity costs of the required resources. Therefore, delaying by one year a project where the initial investment is denoted by I and the stream of benefits and costs goes from 1 to T is profitable if:

$$\frac{rI}{(1+r)} + \frac{B_{T+1} - C_{T+1}}{(1+r)^{T+1}} > \frac{B_1 - C_1}{1+r} ,$$

where the left side of the inequality represents the present value of the benefits associated with postponing the start of the project by one year and the right side is the discounted benefit of the first year, which is lost because postponing the project. If the net benefit of year $T+1$ is small enough, the ‘delay condition’ is straightforward:

²¹ This term is given by $S(T) = \sum_{t=1}^T (1+r)^{-t} = (1/r) [1 - (1+r)^{-T}]$.

$$r > \frac{B_1 - C_1}{I}.$$

Thus, when the discount rate is higher than the rate of return on investment in the first year, ‘waiting’ is the optimal decision.

Postponing a project has an additional benefit, the opportunity of waiting for new information that could change the decision to invest. Infrastructure investment is irreversible and there is usually uncertainty over the future benefits and costs. Moreover, the government can postpone the investment. Under these circumstances there is an opportunity cost of undertaking the investment in the present, since it would imply the loss of the economic value of the information revealed by waiting. To introduce this element in the assessment, the lost option value should be included as a cost of the project. An alternative approach is to compare the *NPV* of two mutually exclusive projects (‘now’ or ‘later’) and choose the one with the highest *NPV*.

In addition to all these technical properties of discounting, the relevance of the social discount rate for the CBA of transport projects also lies in its underlying economic interpretation. When a specific value for the discount rate is selected, the society is also assigning a certain value to the opportunity cost of waiting for the future. Even if a single efficiency-based social discount rate can be agreed upon, there remains plenty of room for discussion on valuation issues, on the effects of capital taxes, on market imperfections or on the uncertainty about the future.

In fact, there are significant variations in social discount rate policies around the world, with developing countries in general applying higher social discount rates (8%-15%) than developed ones (3%-7%), and multilateral investment banks using rates between 10-12%. European Commission recommends a 5% for major projects in Cohesion countries²² and 3% for the other Member States, although different values may be acceptable on the grounds of international macroeconomic trends (see **Annex A**). These variations reflect the different analytical approaches followed by various countries in choosing the rate. But more fundamentally, it can be argued that the divergence reflects differences in the perceived social opportunity cost of public funds across countries.²³

²² The Cohesion Fund is aimed at Member States whose Gross National Income per inhabitant is less than 90% of the EU average. It aims to reduce economic and social disparities and to promote sustainable development.

²³ Recently, there has been interest in using Capital Asset Pricing Model (CAPM) approaches, allowing systematic and project-specific risk.

3.3. The treatment of risk in cost-benefit analysis

3.3.1. Risk sources in project assessment

The economic evaluation of transport projects is seldom performed under certainty conditions. On the contrary, two types of uncertainty must be often dealt with. The first one is project-related uncertainty, which is associated with exogenous contingencies that are yet to occur, affecting the flow of benefits and costs. This happens, for instance, with respect to forecasted demand or the evolution of prices. Once started, the results of the project may be also affected by unforeseen costs or unanticipated institutional or regulatory changes.

The second type of uncertainty is evaluation-related, which refers to the limited information context in which most economic assessments are addressed. Parameters such as the value of time or some elasticities with respect to prices or income are often estimated using external references, but they should not be necessarily considered as deterministic.

Along with exogenous variables, human factors may also cause a project to fall short of expectations: errors are sometimes made in estimating net benefits, in the technical design or even preparing the budget. In many cases, particularly under weak governance and poor institutional design, these errors tend to create a systematic bias in favour of the project: demand is generally overestimated and costs underestimated. It is critical to distinguish whether the source of these systematic deviations are originated by optimism bias or strategic misrepresentation.²⁴

These uncertainties imply that, in practice, most economic evaluations of transport projects are actually performed under risk conditions. Private investors incorporate this circumstance into their decision-making process according to their risk attitude. If they are risk neutrals, their decision can be based on the expected net present value, $E(NPV)$. For every year t , let us define the yearly net value of benefits and costs, $B_t - C_t$, as a random variable that may take different values: $(B_t - C_t)_i$ with $i = 1, \dots, n$. Let us denote by π_i^t the probability that, in year t , the net value of benefits and costs takes value $i: (B_t - C_t)_i$ with $i = 1, \dots, n$. Thus, $\sum_{i=1}^n \pi_i^t = 1$, for every year $t = 0, \dots, T$. The expected net present value is, then, given by:

²⁴ This idea will be expanded in **Section 9**. Further discussion potential biases in decision-making on mega-projects can be found in Flyvbjerg *et al.* (2018).

$$E(NPV) = \sum_{t=0}^T \sum_{i=1}^n \frac{\pi_i^t (B_t - C_t)_i}{(1+r)^t}.$$

In the more frequent case of decision-makers being risk averse, their utility is more negatively affected by the variability of the project results. Although the expected value of net benefits can still be calculated every year, most people in these circumstances are willing to accept a fixed value, lower than the expected one, to avoid the cost of bearing the risk. This is the so-called ‘certainty equivalent’, whose difference with the expected value will be increasing with the investor’s degree of risk aversion. Certainty equivalents can be used in the expression above to obtain the corresponding *NPV*, which would become now the ideal reference to apply the decision criteria. Finding the certainty equivalents is difficult in some contexts, and the treatment of risk in private projects is usually based on adding a risk premium to the interest rate for financial evaluation. Although this approach reduces the project profitability, it is not the same as using certainty equivalents since, for example, risk may affect in different ways benefits and costs.

In the case of the CBA of transport projects, it could be argued in principle that, to avoid distortions in the allocation of resources, their economic evaluation should be carried out using the same risk-adjusted discount rate as in the private sector. However, it is also true that many similar and independent projects are continually being promoted by the public sector and, therefore, by the *Law of large numbers*, their results will collectively tend to revert to their expected values (as in any other insurance mechanism). This would justify adopting the risk neutrality hypothesis and always performing the evaluation of public projects using expected values and risk-free (social) discount rates.

Nevertheless, the above argument of risk pooling is not applicable to singular projects to justify risk neutrality in the public sector. The Arrow-Lind theorem reinforces risk neutrality by appealing to a risk-spreading argument. If the costs of the resources used in the project are equally shared by a large enough number of people (all the taxpayers), the impact on their individual income will be very small: the cost of risk bearing tends to zero and the expected value of net benefits will be close to the certainty equivalent. This is no longer true in large projects that represent a significant contribution for any given social group or when the risk is borne by a small group of people. In these cases, the certainty equivalent should be estimated.

3.3.2. *Dealing with risk in project assessment*

The easiest and most commonly used approach to deal with risk in project assessment is sensitivity analysis in any of its variations. Basically, sensitivity analysis, consists in checking how the *NPV* changes when changing the value of a single critical variable

(e.g. the value of time). This *ad hoc* approach to risk assessment can be conducted variable by variable (sensitivity analysis) or by groups (scenarios).

When several variables are simultaneously modified, different scenarios arise. A typical CBA in this case has, for example, a ‘pessimistic’ scenario (low demand and/or high costs) and an ‘optimistic scenario’ (high demand and/or low costs), which can be compared with a ‘baseline’ or other scenarios by the decision-maker using the standard decision criteria. The main advantage associated with the use of scenarios and, in general, with sensitivity analysis is that it has the ability to reveal the robustness of the results in a very simple and direct way. The main disadvantage is that the procedure is somewhat arbitrary and also ignore the possibility of correlation between variables.

The calculation of switching values is also interesting. It is a variation of the sensitivity analysis based on the calculation of thresholds. The switching value is the value that a critical variable (e.g. demand) would have to take in order for the *NPV* of the project to become zero, or more generally, for the outcome of the project to fall below the minimum level of acceptability.

One procedure in projects subject to very high demand uncertainty is to invert the process in the CBA of the project. Instead of computing the standard *NPV*, the practitioner should calculate the minimum demand in the first year of the project (given one or several growth rates for yearly net benefits) that makes the project socially profitable. This requires solving the equation:

$$NPV(x_{\min}) = 0,$$

which is easy to calculate and produce a reference value (x_{\min}) that can be used to compare with the actual data of population, passengers or freight in the corridor affected by the project.

The best alternative to sensitivity analysis is *risk analysis*, which requires a complete modelling of all the determinants of the *NPV*, which becomes itself a stochastic variable with its own probability distribution. In general, any risk analysis should identify all critical variables and parameters in the model according to the relevance of their impact on the *NPV*, the likelihood of their future variability or the available information to the decision-maker.

In transport projects they typically include overruns in investment costs, demand, time values or other costs. For each of these, a probability distribution should be selected, depending on the existing *a priori* information. For example, if only upper and lower limits are known, a usual choice is the uniform distribution (either discrete or continuous), or triangular (not necessarily symmetry) if we know the most likely value

within the range. If the mean and variance are known (or estimated), another obvious choice is the normal distribution. The analyst may complete the available information by referring to similar projects or using standardised values from official guidelines.

It must be noted, however, that an excessive number of variables complicates the model: the interaction among them must be considered to avoid inconsistent results (e.g., the individuals' wealth is increasing but travel demand decreases). This requires an explicit modelling of the correlation coefficients of the selected critical variables.²⁵ In practice, risk analysis is performed with the help of a specific computer program. Once the project and the variables have been modelled, a large number of simulations can be performed by randomly drawing values according to their probability distribution. The result, as stated before, is a range of (expected) *NPV*s with their respective probabilities.

3.3.3. Decision criteria using risk analysis

Risk analysis does not only provide a very useful quantitative information in terms of the net present value of the benefits and costs of the project; it also gives the probability of the *NPV* being positive or negative. At least two reasons justify the relevance of this latter information. First, when the decision-maker is risk averse and wants to know beyond the sign of the expected *NPV*. A decision-maker facing a project with a positive *NPV* may be interested in the information provided by a risk analysis. A project with a positive *NPV* and a 1% probability of a negative outcome is not the same as a project with the same expected *NPV* and a 40% probability of yielding a negative *NPV* (the one represented on the left-hand side of **Figure 3.4**). Second, when the project represents an investment of such magnitude that its distribution among the taxpayers makes their individual contributions significant. Since the Arrow-Lind theorem does not hold in this case, it is necessary to explicitly consider the cost of risk bearing.

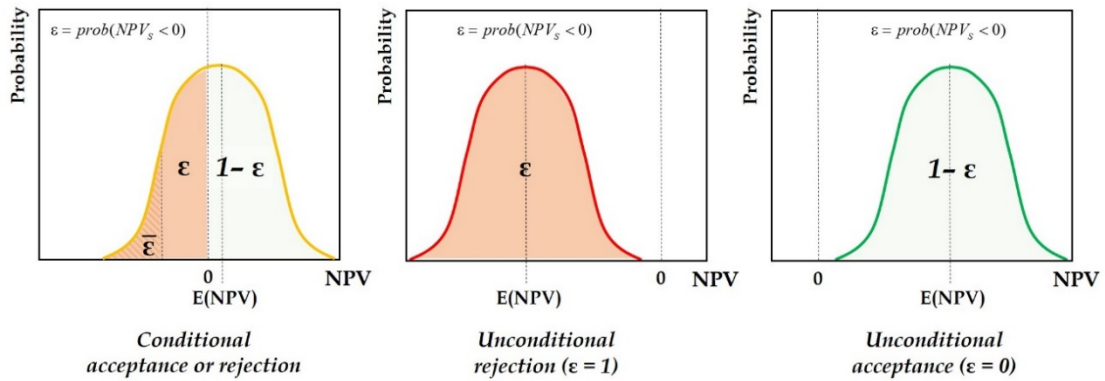
The information provided by risk analysis makes it also possible to reformulate the decision criteria in probabilistic terms. Consider first that a single project is under assessment, and it must be decided whether to carry it out or to reject (or delay) it. Since there is not a unique value for the social and financial net present values, the decision must rely, in principle, on their complete probability distribution. It is possible however to focus only on the critical probability associated with negative outcomes,

²⁵ Although the precise nature of such correlations is often unknown and cannot be specified with a great deal of accuracy, the problem cannot be ignored, since it may distort the results of risk analysis. The reason for this is that the selection of input values from the assigned probability distributions for each variable is purely random. It is therefore possible that the resultant inputs generated for some scenarios violate a systematic relationship that may exist between two or more variables (Savvides, 1994).

defined as $\varepsilon = \text{prob}(NPV < 0)$, and simultaneously assume that there exists an exogenous reference probability, $\bar{\varepsilon}$, above which the decision-maker is not prepared to accept the project. This reference value must be implicitly or explicitly revealed, and will be lower the higher the degree of risk-aversion of the decision-maker.²⁶ Thus, the general decision criterion in this case is very simple and will be always conditioned to the risk-aversion degree of the decision-maker: the project must be (conditionally) rejected if $\varepsilon > \bar{\varepsilon}$, and (conditionally) accepted if $\varepsilon \leq \bar{\varepsilon}$.

This idea is illustrated in the left panel of **Figure 3.4**, where the probability density function of the social net present value obtained from risk analysis is represented, and the areas (probabilities) ε and $1 - \varepsilon$ are correspondingly defined. If the probability of the social net present value being negative is large enough as compared to an exogenous reference, $\bar{\varepsilon}$ (as represented) the project must be rejected (and vice versa). The middle and right panels in the figure illustrate two extreme cases: the project must be unconditionally rejected if $\varepsilon = 1$ (since it will always yield negative results in social terms) and unconditionally accepted if $\varepsilon = 0$ (it will always yield positive results). In any case, if the decision-maker faces budget constraints, projects that have been (conditionally or unconditionally) accepted must additionally be analysed in terms of their financial viability.

Figure 3.4. Decision criteria under uncertainty: accept/reject decisions



When comparing two or more projects, they should also be ranked according to the probability density functions of their social and financial net present values. In most cases, however, the comparison can rely on two single parameters: the expected value and the variance (as a measure of the risk associated with the variability of the project results). In general, projects with higher expected values and lower variances should

²⁶ For risk-neutral individuals, decisions will be made using the $E(NPV) > 0$ criterion, as stated above. Note also that the reference value may also differ when considering social and financial NPV .

be chosen first, but there may be situations where the decision will not be so uncontroversial.

4. PRICING AND INVESTMENT DECISIONS IN TRANSPORT PROJECTS

4.1. Pricing and investment when there are no other transport alternatives

Making decisions about accepting or rejecting projects, or about their ranking when there are different alternatives, always requires taking into account how external interventions in transport markets affect the successive equilibria that occur in these and other related markets. This is hardly compatible with the assumption that prices are completely exogenous and, therefore, we have to be aware that the social profitability of any given project may differ under different pricing policies. Prices affect demand and, thus, the size (or even the sign) of agents' surpluses and whether the revenues will be enough to cover costs. Therefore, when deciding whether it is socially desirable to invest in a project or not, its economic evaluation should consider the future charging scheme (including the cases where no prices are charged at all) and whether the project requires the use of public funds.

Let us consider as an example a transport project consisting in building a new transport infrastructure under the assumption that there are no other transport alternatives to address the same (mobility) problem. We will first assume that there is no economic cost of public funds, which is defined as the loss incurred by society when raising revenues to finance government spending through distortionary taxation. Secondly, we will discuss the role of the pricing scheme and the economic cost of public funds in the investment decision. In all cases, we will assume that there is enough capacity to satisfy the demand, that is, there are not congestion problems (see **Section 6**).

4.1.1. Investment decisions when there is no economic cost of public funds

Consider the evaluation of a new transport infrastructure that requires an initial investment denoted by I . In order to illustrate the importance of the pricing scheme on this investment decision, throughout this subsection we will define the change in social welfare due to the project as the sum of consumers and producers surpluses net of the investment cost, $\Delta\bar{W}$.

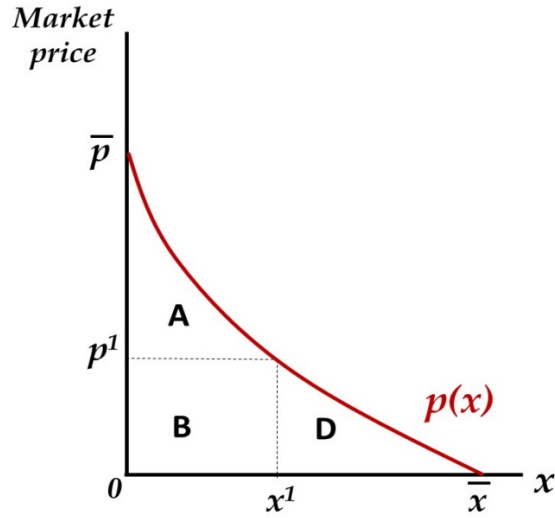
Thus, according to the criteria defined in **Section 3**, the new infrastructure should be built if the net increase in social welfare, $\Delta\bar{W}$, is higher than the investment cost, I :²⁷

²⁷ This simplified approach follows the classical CBA example by Jules Dupuit (as reprinted in Dupuit, 1995), regarding the social welfare effects of building a toll-free bridge.

$$\Delta \bar{W} = \Delta CS + \Delta PS > I. \quad (4.1)$$

Let us start assuming that there are no operating costs, there are no other transport alternatives, and there is no economic cost of public funds. **Figure 4.1** represents the inverse demand function for the use of such a new transport infrastructure.

Figure 4.1. Pricing and social welfare when there are no operating costs



The increase in social welfare due to this transport project depends on the pricing scheme. If there is free access to the transport infrastructure ($p = 0$), the net change in social welfare is equal to the increase in consumers' surplus due to the project (given by **area A + B + D**), since producers' surplus is equal to zero. Although, in this case, free access is the optimal pricing scheme, the government may be budget-constrained and a positive price for the use of the new infrastructure may be needed.

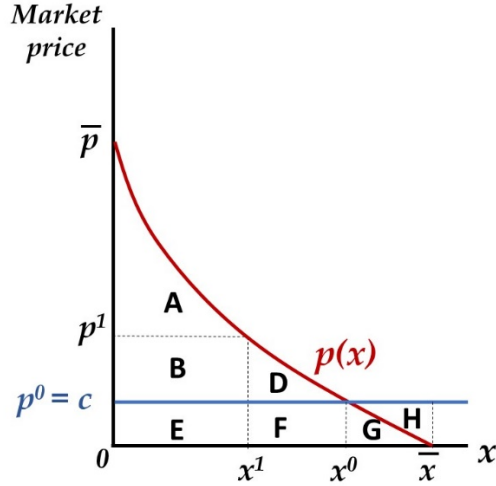
If this price is equal to p^1 , the net change in social welfare is given by the increase in consumers' surplus (**area A**) and the increase in producers' surplus (**area B**). Therefore, the social profitability of the project critically depends on the pricing scheme, since:

$$\begin{aligned} \Delta \bar{W}_{p=0} &= A + B + D, \\ \Delta \bar{W}_{p=p^1} &= A + B. \end{aligned}$$

How does the charging scheme affect the investment decision? If $A + B > I$, the optimal decision is to accept the project, regardless on whether the government plans to charge p^1 or allow free access. On the contrary, if $A + B + D < I$, the optimal decision is not to invest independently on the charging scheme. Interestingly, it might be the case that $A + B + D > I > A + B$, that is, under a free access scheme, the infrastructure

should be constructed, but if the government is planning to charge p^1 , the socially optimal decision is not to invest in this transport infrastructure. Therefore, the investment decision may strongly depend on the pricing scheme to be used once the infrastructure has been constructed.

Figure 4.2. Pricing and social welfare with constant marginal operating costs



Now consider that marginal operating costs are constant and equal to $c > 0$, as in **Figure 4.2**. If there is free access, the net change in social welfare is equal to the increase in consumers' surplus, given by areas **A + B + D + E + F + G** minus the decrease in producers' surplus, given by **E + F + G + H**. Thus, the net change in social welfare is **A + B + D - H**. Conversely, if the government applies an optimal marginal cost pricing scheme ($p^0 = c$), the net change in social welfare is equal only to the increase in consumers' surplus (**area A + B + D**), since, in this case, producers' surplus is equal to zero. Finally, if the price is set above marginal costs, $p^1 > p^0$, the net change in social welfare due to the transport project is given by the increase in consumers' surplus (**area A**), and the increase in producers' surplus (**area B**). These results can be summarized as:

$$\Delta \bar{W}_{p=0} = A + B + D - H,$$

$$\Delta \bar{W}_{p=c} = A + B + D,$$

$$\Delta \bar{W}_{p=p^1} = A + B,$$

showing that, once again, the optimal investment decision according to expression **(4.1)** may strongly depend on the pricing scheme. Suppose, for example, that the values in the figure are such that **A + B + D > I**, but **A + B + D - H < I** and **A + B < I**. In this case, it is socially optimal to invest in the construction of the new infrastructure if the government is planning to charge a price equal to marginal

operating costs, but the project should be rejected if free access or prices higher than marginal operating costs are being considered as pricing policies.

4.1.2. Investment decisions when there is an economic cost of public funds

Let us now consider that public funds are obtained through distortionary taxation and, thus, there is a shadow price of public funds denoted by $\lambda_g > 1$. This means that when the government uses I monetary units of public funds to build the new transport infrastructure, this investment really costs $\lambda_g I$ to the society in terms of the distortions created in the economy when raising taxes to finance such an investment.²⁸ However, if the government charges a positive price for the use of the new infrastructure and collects net revenues equal to $(p - c)x$, this amount can be used to reduce existing distortionary taxation and, thus, the real profit for the society is $\lambda_g (p - c)x$.

Therefore, defining the net increase in social welfare due to the transport project as the sum of consumers' surplus and producers' surplus (net of the investment cost), the 'acceptability condition' for the new infrastructure becomes:

$$\Delta \overline{W} = \Delta CS + \lambda_g \Delta PS > \lambda_g I. \quad (4.2)$$

Again, let us start the discussion assuming that there are no operating costs, as in **Figure 4.1**. If there is free access to the transport infrastructure, the only net change in social welfare due to the project is equal to the increase in consumers' surplus (**A** + **B** + **D**), since producers' surplus is equal to zero. On the contrary, if the government charges a price p^1 for the use of the infrastructure, the net change in social welfare is given by the increase in consumers' surplus (**area A**) plus the increase in producers' surplus (**area B** multiplied by the shadow price of public funds). Once more, the social profitability of the project depends on the pricing scheme:

$$\begin{aligned} \Delta \overline{W}_{p=0} &= A + B + D, \\ \Delta \overline{W}_{p=p^1} &= A + \lambda_g B. \end{aligned}$$

How does the shadow price of public funds affect the investment decision? If λ_g is high enough it might be that $A + \lambda_g B > \lambda_g I > A + B + D$. In this case, if there is free access to the transport infrastructure, the socially optimal decision is to reject the project, but if the government is planning to charge a price p^1 , the socially optimal decision is to build the infrastructure.

²⁸ For further details about the marginal cost of public funds see Gahvari (2006) and Dahlby (2008).

Suppose now that marginal operating costs are constant and equal to c as in **Figure 4.2**. Under a free access pricing scheme, the net change in social welfare is again equal to the increase in consumers' surplus ($A + B + D + E + F + G$) minus the decrease in producers' surplus, given by $\lambda_g(E + F + G + H)$. If the price is $p^0 = c$, the net change in social welfare is just equal to the increase in consumers' surplus ($A + B + D$), because producers' surplus is equal to zero. Finally, if the government charges p^1 , the net change in social welfare is given by the increase in consumers' surplus (A) and the increase in producers' surplus ($\lambda_g B$). In sum, we have that:

$$\begin{aligned}\overline{\Delta W}_{p=0} &= A + B + D + (1 - \lambda_g)(E + F + G) - \lambda_g H, \\ \overline{\Delta W}_{p=c} &= A + B + D, \\ \overline{\Delta W}_{p=p^1} &= A + \lambda_g B.\end{aligned}$$

If the shadow price of public funds is large enough, we could have, for example, that $A + \lambda_g B > \lambda_g I > A + B + D$. In this case, it would be socially optimal to invest in the construction of the new infrastructure if the government is planning to charge prices above marginal operating costs, but the transport infrastructure should not be built if the government is planning to allow free access or charge a price equal to marginal operating costs. So, once again, we have shown that the investment decision may strongly depend on the pricing scheme.

4.2. Pricing and investment when there are several transport alternatives

We will now discuss the relationship between pricing and investment when deciding on alternative transport infrastructures. Pricing decisions affect differently the social welfare of alternative transport modes. Thus, when comparing different transport alternatives, a particular charging scheme may favour the creation of a particular transport infrastructure network, leading to long-term equilibria that would not be optimal under other charging schemes. This relationship between decision-making and pricing stresses the relevance of project planning and the link between evaluation and pricing.²⁹

In order to illustrate this idea, consider two regions connected through air transport. For these two regions the government must decide whether to invest in a new high-speed rail (HSR) infrastructure, using two possible charging schemes: either charging according to short-run marginal costs or charging mark-ups over short-run marginal costs. Although charging according to short-run marginal cost is the general rule in transport infrastructures with sunk investment costs, departures from it may be

²⁹ This subsection draws on de Rus and Socorro (2019).

justified for three reasons (Laffont and Tirole, 1993): under the existence of an economic cost of public funds (shadow price of public funds higher than one); when users do not reveal their willingness to pay for capacity with a short-run marginal cost pricing scheme, and when charging according to short-run marginal costs reduces the incentive for cost reduction. Finally, there might be also competition reasons: when users pay for the full cost of each transport mode, intramodal competition is not affected.³⁰

In this context, we show that the government may favour the construction of the HSR infrastructure by choosing a charging scheme based on short-run marginal costs. In order to optimally decide whether to invest or not in the new infrastructure, the government should compare the social welfare of this project with the social welfare obtained in the situation in which the regions are only connected through air transport. Social welfare is defined as the sum of users' surplus, transport operators' surplus, and profits due to the use of the transport infrastructure, minus the opportunity cost of the investment in the HSR infrastructure. Assuming for simplicity, that yearly net social benefits are constant and equal to \bar{B} , and the HSR lasts forever, the social net present value of introducing the HSR can be written as:

$$NPV_s = -I + \frac{\bar{B}}{r},$$

where I is the investment cost in year zero, \bar{B} represents annual net benefits starting in year 1, constant during the whole life of the project, and r is the social discount rate. In this case, $NPV_s > 0$ implies that $\bar{B} > rI$, where rI is the opportunity cost of the investment per year.

Under these assumptions, a positive NPV_s is equivalent to fulfil the condition of optimal timing (it is optimal to invest today instead of postponing the investment) as expressed in the following expression:

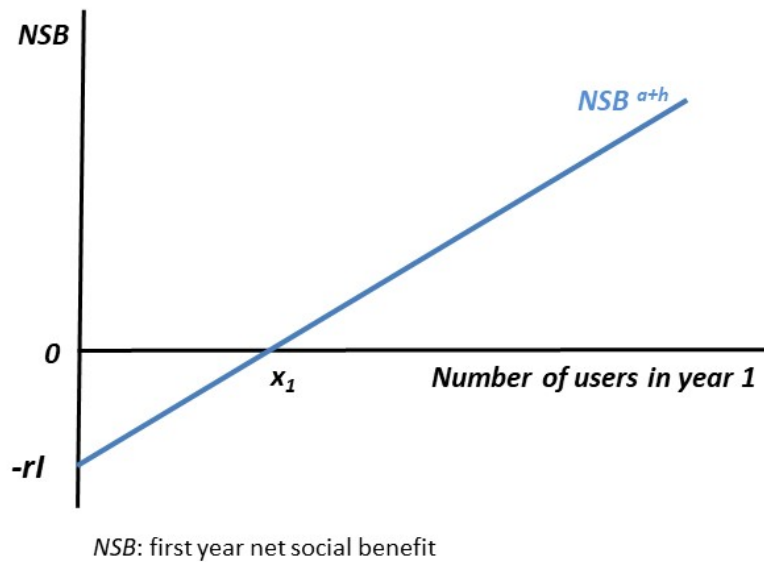
$$\frac{B_1}{1+r} > \frac{rI}{1+r},$$

where B_1 represents first year net benefits. In other words, if net benefits during the first year are higher than the opportunity cost of the investment per year, we should construct the infrastructure today.

³⁰ Equity issues could also be added in favour to the departure from short-run marginal cost pricing.

Figure 4.3 represents the optimal timing condition. The first year net social benefit (NSB) associated with the project of investing in a new HSR infrastructure is represented in the vertical axis, taking into account that regions are already connected through air transport. Notice that the NSB associated with this alternative is strictly increasing in the number of users, x . When $x = 0$ the NSB of the first year associated with the project is just the opportunity cost of the investment of the new transport infrastructure ($-rI$). **Figure 4.3** depicts a situation in which if the number of users in the economy is lower than x_1 the optimal decision is not to invest in the HSR infrastructure and wait till the number of users increases. On the contrary, if the number of users is higher than x_1 , the optimal decision is to invest in the new transport infrastructure and connect the regions through the air transport and HSR.

Figure 4.3. Optimal decision depending on the number of users



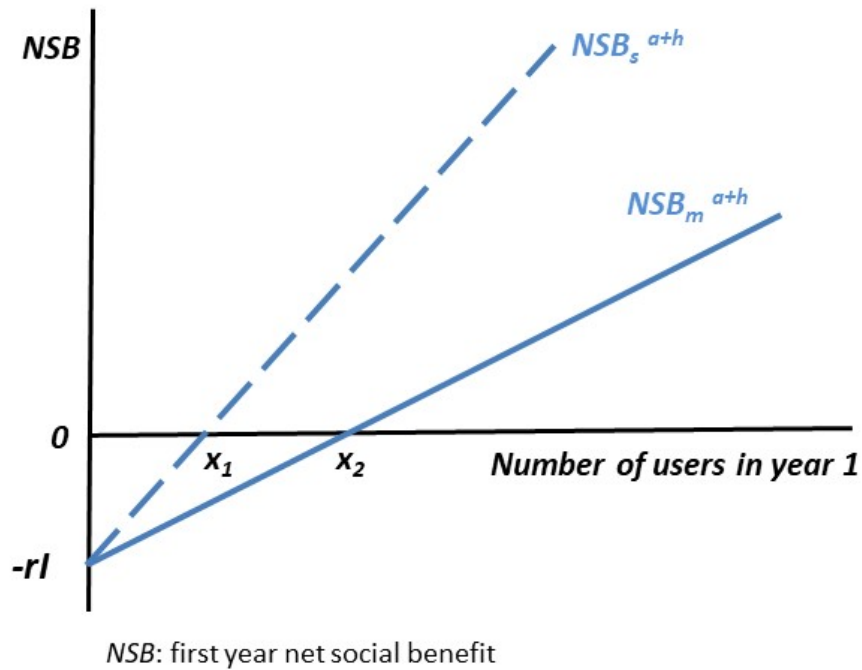
The slope of the NSB associated with the project of constructing the new HSR infrastructure strongly depends on the charging scheme chosen by the government. In particular, the higher the access prices are, the flatter is the NSB of the project. Thus, if the government moves from a charging scheme based on short-run marginal costs to a charging scheme based on mark-ups over short-run marginal costs, the NSB function becomes flatter.

This is illustrated in **Figure 4.4**, where the dashed line represents the NSB of the project when using a charging scheme according to short-run marginal costs, NSB_s^{a+h} , and the solid line represents the NSB function when moving to a charging scheme based on mark-ups over short-run marginal costs, NSB_m^{a+h} . Notice that, when the government decides to charge the HSR infrastructure access using mark-ups over short-run

marginal costs instead of a pricing scheme based on short-run marginal costs (as in **Figure 4.4**), critical thresholds for the number of users move from x_1 to x_2 . Thus, with a charging scheme based on mark-ups over short-run marginal cost, we need more users to make the decision of constructing the new HSR infrastructure optimal.

Notice that if the number of users is between x_1 and x_2 , the socially optimal decision is to have just airports if the government uses a charging scheme based on mark-ups over short-run marginal cost. Nevertheless, if the government charges according to short-run marginal costs the optimal decision is to invest in the new HSR infrastructure (and having both the air and rail infrastructure). Therefore, the government may favour the construction of the HSR infrastructure choosing a charging scheme based on short-run marginal costs.

Figure 4.4 Optimal transport alternative depending on the charging scheme



The practical consequence of this analysis for transport policy is straightforward: Any transport infrastructure should be constructed only in those cases in which the level of social welfare is clearly higher than the social welfare associated with the next best alternative, and this fact strongly depends on the charging scheme. It is worth highlighting that the long-term consequences of investing in suboptimal infrastructure projects can be paramount. Due to the irreversibility of the investment, we may end up with a transport network optimal under a particular charging scheme but suboptimal under any other charging scheme. Once the infrastructure has been constructed, it

should be used (if at least variable costs are covered). However, this does not mean that new segments should be added to the existing network. For the new segments, the planner should carefully choose the charging scheme and wait till the demand reaches the required threshold for social profitability. Meanwhile, postponing the expansion of the network is socially worthy.

There are cases in which two transport networks have been constructed to solve the same mobility problem but using different pricing schemes. Airports and HSR networks are a clear example. The effects on modal split have been very favourable for the HSR, having captured in some lines practically all demand. However, this shift in demand may be due to the fact that HSR users do not pay for the construction cost of the rail infrastructure while in large airports (that is, those competing with HSR) air transport users do.

The discussion on what prices should be charged for the use of transport infrastructures and services remains one of the most controversial in the literature and is still far from being resolved. Although the general rule from the point of view of economic efficiency is that prices should be equal to social marginal costs, these costs can be defined in the short or long-term. In addition, there might be budget constraints and incentive problems (that will be analysed in **Section 9**). In general, the optimal charging scheme of a given transport infrastructure should be made taking into account the existence of other complementary or substitute modes of transport and not only the specific characteristics of each mode (de Rus and Socorro, 2014). This again is a reminder of the importance of a sound transport policy.

5. THE OPPORTUNITY COSTS OF PROJECTS: MARKET AND SHADOW PRICES

5.1. The cost of the resources involved in a project

The social benefits achieved through the reduction of the generalized cost of transport are not free for the society. These benefits always have an opportunity cost that is measured by the value of resources diverted from other uses due to their use in any particular project. This section deals with the inputs needed for the transport project and how to value them.³¹

According to **Section 2**, the effects of a transport project on social welfare can be expressed as the maximum income the affected individuals is willing to pay to enjoy the corresponding benefits, net of the project costs. This is the value of the sum of the compensating variations (CV) in expression (2.12) or (2.20) for all the individuals of the society, which is net of project costs. The aggregation of the CV is then the sum of the individuals' willingness to pay for the benefits of the project (positive sign) and the willingness to accept for giving up other goods to achieve those benefits (negative sign).

Thus, the social opportunity cost of the project (C_j) can be defined as the value of all the goods (s) the society has to give up, when those resources are deviated from other uses in order to implement the project, that is, to enjoy its benefits on good j (e.g. a faster transport service), as formally represented by:

$$C_j = \sum_{k=1}^s p_k dx_k, \quad (5.1)$$

an expression derived from the general model in **Section 2**, with $s \leq n$ goods or services, and where it is assumed that the only input, labour, is fully utilized to produce and consume goods and that the market prices reflects the value of the goods deviated to the project.

The problem is that this expression is not very useful for computing the NPV of the project. The CBA practitioner generally does not know which goods (schools, housing, leisure facilities, etc.) are given up by the society to achieve the benefits of the project under assessment. However, there is a way to circumvent this measurement problem using the same approach suggested by (5.1). To do so, we can now consider that the

³¹ This section draws on Johansson (1993) and de Rus (2010). In particular, the section deals with inputs that can be purchased in markets. Non-market resources are discussed in **Section 6**.

production of any of the goods in this expression, x_k , requires at least two inputs y_1 and y_2 , and the corresponding production function is then given by

$$x_k = f_k(y_1, y_2), \quad (5.2)$$

whose total differential shows that any output variation depends on the change in the quantity of the inputs used multiplied by their marginal productivities:

$$dx_k = \frac{\partial f_k}{\partial y_1} dy_1 + \frac{\partial f_k}{\partial y_2} dy_2. \quad (5.3)$$

Replacing (5.3) in (5.1) and recalling that any profit maximizing firm uses additional units of inputs until their market price (w) equals the value of its marginal productivity, $w = p(\partial x / \partial y)$, the cost of the project can be re-expressed as:

$$C_j = \sum_{k=1}^s (w_1 dy_1 + w_2 dy_2). \quad (5.4)$$

The cost of the project initially expressed in (5.1) as the social value of the diverted goods, to get the good provided by the project appears now in (5.4) as the quantities of the inputs (dy_1 and dy_2) required for the production of those goods, multiplied by their respective prices (w_1 and w_2).

In practice, the validity and usefulness of expression (5.4) for identifying and assessing the costs of a project is conditioned by three underlying assumptions. Firstly, all the changes in input markets are marginal; secondly, input and output markets are perfectly competitive, without distortions, like indirect or income taxes; and thirdly, all the resources are fully utilised. However, once these assumptions are abandoned to deal with more realistic project assessment situations (that include, among others, the presence of subsidies or taxes, or the use of unemployed labour in the project), expression (5.4) is no longer valid to calculate the opportunity costs of the project. This is what shadow pricing is about: adjusting market prices to reflect the opportunity costs.

5.2. Market and shadow price of inputs

A transport project typically requires the use of some produced goods (for example, vehicles, x) and inputs (y , such as energy, spare parts and other materials), from which, for convenience, we particularly separate, labour (denoted as L , in terms of number of workers). Thus, as showed by (5.5) the social cost of this project (C) can be rewritten as:

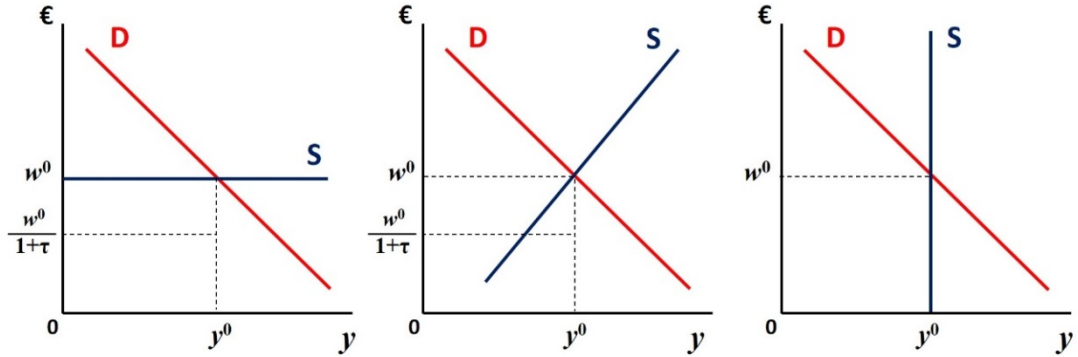
$$C = p dx + w_y dy + w dL, \quad (5.5)$$

where, p is the vector of market prices (including taxes) of the goods used by the project, w_y the vector of market prices of inputs other than labour, and w the wage rate. For example, a high-speed rail project requires building a dedicated infrastructure (tracks and stations), producing all the equipment required for this infrastructure and the operation of the services (such as construction materials, rolling stock, etc.), and hiring labour force for the maintenance and operation of the infrastructure and services.

In expression (5.5), the cost of the locomotives, wagons, electric power and labour, are initially valued through market prices. We now discuss when this is correct and when some adjustments in those prices should be made to obtain the opportunity costs of the goods, inputs and labour. Note also that the distinction between goods and inputs in (5.5) is somehow blurred in practice as the inputs to be purchased for the project, are indeed produced inputs (i.e., goods). Nevertheless, we keep the distinction for a later discussion of the shadow price of inputs, where those inputs are deviated from the private sector.

Figure 5.1 illustrates these ideas by showing the supply and demand in equilibrium of any input y , where (w^0, y^0) represents the market equilibrium under different supply schemes and there is an *ad-valorem* indirect tax (τ). Whereas the demand function is the same for the three cases, the supply is perfectly elastic in the left panel, showing the (infinite) availability of any quantity at the market price w^0 . In the central panel, the input supply function is upward-sloping, indicating that to supply more than y^0 requires an increase in w . In the right panel, the supply is perfectly inelastic meaning that the available amount of the input is fixed at y^0 and cannot be increased.

Figure 5.1. Market and shadow prices of inputs



Now consider that the project (marginally) shifts the input demand to the right. In the case of a perfectly elastic supply, the input price w^0 remains constant and the suppliers provide the additional amount of y , that is, dy in expression (5.5). Note that w^0 is not the opportunity cost of an additional unit of y , as the indirect tax (τ) is a mere transfer. Therefore, the shadow price of y for our project is the market price of the input net of taxes: $w^0/(1+\tau)$.

In general, the shadow price of any of the inputs used in transport projects are the market price of the input, net of taxes. This principle is applicable to the majority of projects whose demand is marginal in the input markets, even for the central case of **Figure 5.1** if the amount of inputs demanded by the project is marginal with respect to national or international markets. In all these cases, the demand shift is negligible, and the equilibrium price w^0 only changes marginally.

Under this assumption, the central panel of **Figure 5.1** is also useful to go deeper in the identification of the opportunity cost of the project. The quantity of y demanded for the project shifts outwards the demand curve in the amount dy , and the input price w^0 goes up to clear the market. This increase in w induces a marginal increase in the supply of y (to the right of y^0) and also a deviation of y from the private sector in the purchase of y (to the left of y^0) when some firms find no longer profitable to purchase y^0 at a higher price than w^0 .

The opportunity cost of the quantity of the input y purchased for the project has a lower and an upper bound depending on where that amount of dy comes from. When one unit of y comes from the additional production of y , the opportunity cost is the marginal social cost of producing the input, i.e., the market price of the input net of the indirect tax, as before, $w^0/(1+\tau)$. Alternatively, when one unit of y is deviated from the production of other goods in the private sector, the opportunity cost of one unit of y is

the value of the production associated with that input in the private sector. In this case, the marginal social cost of y is higher, and equal to w^0 in the central panel of the figure.

The opportunity cost of the input deviated from the private sector to the project could be even higher. In practice, if the private sector is paying an ad-valorem tax (θ) for the production of a good crowded out by the diverted input, the opportunity cost of one unit of y is $w(1 + \theta)$, where θ is the added value tax. Thus, $w(1 + \theta)$ is equal to the value of the marginal productivity of the input.

An example clarifies this idea: if the marginal cost of one litre of petrol is €0.5 and there is a specific revenue collecting tax on petrol of €0.5, its market price (w) is €1 per litre. The social opportunity cost of using an additional litre of petrol in the project is 0.5, when the litre comes from new production (as in the left panel of **Figure 5.1**), and €1.2 when the VAT is 20% and the litre of petrol is deviated from a private transport company. Hence, lower bound of the shadow price of one litre of petrol is €0.5, and the upper bound €1.2.

In the case of produced goods, the market price (p) in expression (5.1) is the opportunity cost. When the produced good is additional, the opportunity cost is lower because the *ad-valorem* tax and any other specific taxes should be deducted to get the social marginal cost.

In practice, inputs are generally available, as in the left panel of **Figure 5.1**, and their shadow price is, simply, its market price net of the indirect taxes. In the case of a shortage or rationing of the input, the rules explained above apply. The distinction between the produced good and the input is now clearer.

Finally, note that the derivation of the rules to convert market into shadow prices were obtained with indirect taxes or commodity specific taxes designed to raise revenue. In the case of Pigouvian taxes, i.e., taxes to internalize externalities, the opportunity cost includes the tax, as the tax is reflecting an external cost and not a mere income transfer as in the case of VAT and similar taxes (provided that the shadow price of public funds is equal to one).

The case of inputs that are available in fixed supply is particularly relevant in the case of transport infrastructure. A project such as the construction of an airport or a railway line requires land. The payment of this fixed factor is called economic rent, meaning that the payment of the factor is above the minimum price to have the factor offered in the market. This is the case of land, represented in the right panel of **Figure 5.1**, with a fixed supply in the long-run and with an equilibrium price w^0 determined by the value of the demand of the factors. This means that the price of this input depends on the value of the products produced and sold using that land (agriculture, commercial or any other economic activity). The point is that for a private firm the land is a cost but

for the society the value of the land is simply a reflection of the value of the goods obtained using the land.

The cost of the land for the project is based on a simple idea. If the land market is competitive, the opportunity cost of land required for a transport project is its market price (w^0), reflecting the net benefit lost in the best possible alternative use of that land. When the project increases the demand of land, the equilibrium represented in the right panel of **Figure 5.1**, is no longer compatible with the initial price w^0 . The quantity of land demanded by the project shifts outwards the demand curve in the amount dy , and the price of land w^0 goes up to clear the market. When the price of land rises, the quantity of land demanded by the private sector goes down, releasing the land required by the project. Hence, the opportunity cost of the land for the project is the value lost when the project displaces the economic activity in that piece of land, and for small quantities of land it can be approximate by the market price of land.

In those cases where the project represents a significant change in the demand for land, the outward shift of the demand curve (in the amount of land required by the project) will increase the price, from w^0 to w^1 and the amount of land required for the project can be calculated with the average of both prices, $(w^0 + w^1)/2$, multiplied by dy . Finally, it should be underlined that the social opportunity cost of land is rarely the payment for the expropriation which are based on some official values non necessarily equal to the willingness to accept of the individuals to release voluntarily this land for the project.

5.3. The opportunity cost of labour

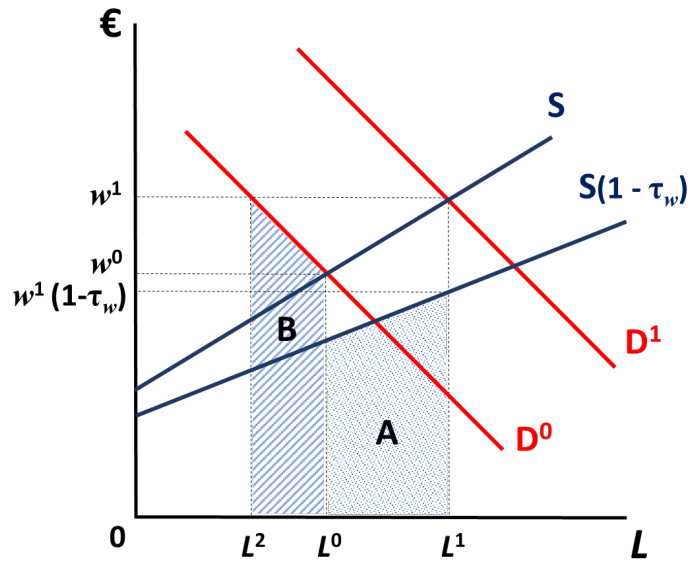
The analysis of the shadow price of labour is virtually the same as the one applied to other inputs. Nevertheless, there are some differences that deserve a separate treatment of this input. Labour is required in the design and construction of transport infrastructure, in its maintenance and operation and in the provision of transport services using that infrastructure.

The opportunity cost of labour in expression (5.5) is valued at its market price, w , but again this is only valid under several restrictive assumptions that usually do not hold in real project assessments, in particular regarding the existence of unemployment. Thus, once the number of workers required for the project is known, the next step is to identify where these workers come from. Were they already working in the private sector? Were they previously unemployed? Were they receiving unemployment benefits? In the analysis of the shadow price of labour it is advisable to distinguish three main possible sources of the labour demanded by a project: (a) workers already

employed in other productive activities; (b) voluntarily unemployed at the current wage; and (c) involuntarily unemployed, willing to work at the current wage.

Figure 5.2 illustrates the necessary adjustments to go from the labour market wage rate to the opportunity cost of labour in each of these cases. We will assume that the project will have a significant effect on the demand of labour and that there is a proportional income tax, τ_w .

Figure 5.2. The shadow price of labour



Initially, without the project, the labour market is in equilibrium with the supply (S) and the demand (D^0) determining a wage rate of w^0 and a quantity of labour of L^0 . The existence of an proportional income tax (τ_w) introduces a distinction between the market supply function (S) and the opportunity cost of the labour supplier, $S(1 - \tau_w)$. The supply function $S(1 - \tau_w)$ shows the marginal value of leisure to the workers and the demand function the value of the marginal productivity of labour for the firm. At the equilibrium wage rate (w^0), the value of marginal productivity of labour for the firm is equal to the value of leisure for the marginal worker plus the income tax.

With the project, the demand of labour shifts from D^0 to D^1 (the horizontal distance between these two parallel demands is exactly the amount of labour required for the project), the wage rate goes up to w^1 and the private demand for labour goes down from L^0 to L^2 . The increase in the wage rate has also the effect of increasing the number of workers willing to work at this higher wage rate, and the equilibrium number of workers goes up to L^1 . Now, we are ready to calculate the opportunity cost of the labour employed in the project.

The project needs $(L^1 - L^2)$ units of labour. This quantity of labour has two components: new workers who wants to work at the new equilibrium wage ($L^1 - L^0$)

and workers already employed in the private sector, $(L^0 - L^2)$, who shift to the project at the higher wage w^1 . The opportunity cost of previously voluntarily unemployed workers, i.e., $(L^1 - L^0)$, is represented by **area A**, the value of leisure lost when they accept the new jobs. Although they are paid $w^1(1 - \tau_w)(L^1 - L^0)$, their social opportunity cost is:

$$\left[\frac{1}{2}(w^0 + w^1)(1 - \tau_w) \right] (L^1 - L^0).$$

The opportunity cost of those already working in the private sector $(L^0 - L^2)$ who shift to the project at the higher wage w^1 , are also paid $w^1(1 - \tau_w)$. However, the social opportunity cost of these workers is higher, represented by **area B**, the lost value of the marginal productivity of labour in the private sector, when the amount of labour $(L^0 - L^2)$ shifts to the project. They are paid $w^1(1 - \tau_w)(L^0 - L^2)$ but the social opportunity costs of these workers is, in principle:

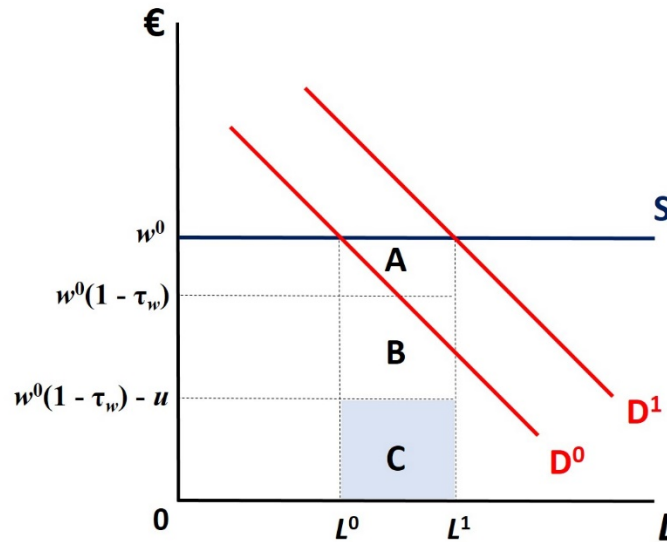
$$\frac{1}{2}(w^0 + w^1)(L^0 - L^2).$$

This is the opportunity cost of the deviated labour when w the unit cost of labour for the firm as represented in **Figure 5.2**. In the case of a proportional social security contribution paid by employers (α_w) plus the existence of *ad-valorem* indirect taxes (e.g. VAT), levied on the product market, the shadow price of the deviated labour has to reflect the social value lost as a consequence of displacing labour from other productive activities. This includes the tax revenues and any other charges lost in the process. The shadow price of labour is in this later case:

$$(1 + \theta)(1 + \alpha_w) \left[\frac{1}{2}(w^0 + w^1)(L^0 - L^2) \right].$$

When the labour of the project involuntarily unemployed, willing to work at the current wage, the supply curve is perfectly elastic, as in **Figure 5.3**.

Figure 5.3. The shadow price of labour with involuntary unemployment



The supply has an infinite elasticity showing that the workers are willing to work at the equilibrium wage if they are hired by the firms. At the level of demand, D^0 there is involuntary unemployment. The project shifts the demand of labour from D^0 to D^1 . The project requires $(L^1 - L^0)$ units of labour, the distance between the demand without and with the project, and this amount is supplied to the market without any change in the initial wage rate.

Here it is useful to distinguish between the worker opportunity cost and the social opportunity cost. **Figure 5.3** shows that the unemployed worker receives unemployment benefits equal to u , and if he accepts the job there is a proportional income tax (τ_w). Hence, as his reservation wage is w^0 (he is not willing to work for less than this wage), the workers' payment is equal to the value of leisure plus the unemployment benefits (u) plus the income tax (τ_w) he has to pay if he accepts the job.

The individual opportunity cost is, therefore, his value of leisure plus the unemployment benefits. Nevertheless, the social opportunity cost cannot include the unemployment benefits (a transfer) as a cost of the project because the real lost in resources when the individual is employed is only the marginal value of leisure. The

shadow price of labour is then $w^0(1 - \tau_w) - u$, and the social cost of these workers for the project is $[w^0(1 - \tau_w) - u](L^1 - L^0)$ corresponding to **area C** in the figure.³²

Both values are useful in the economic evaluation of projects. The point is to be consistent with the alternative approaches the analyst can follow. In the case of adding the change in surpluses, the private opportunity cost is what matters, whereas it is the social opportunity cost that is relevant if the approach is the change in willingness to pay and resources.

The social opportunity cost of $(L^1 - L^0)$ in **Figure 5.3** is represented by **area C**, the value of leisure. The private opportunity cost is higher and represented by the **areas B+C**. Adding the change in surpluses, the change in producer surplus is zero, as they receive **A+B+C** (the value of the marginal productivity of labour) and they pay **A+B+C**. Workers' surplus is also zero, and they receive a gross payment of **A+B+C**, equal to their opportunity cost. Finally, taxpayers receive the income tax (**area A**) and the savings in unemployment benefits (**area B**). Adding these surpluses, we get **A+B** as the net benefit, resulting from the value of the marginal productivity of labour (**A+B+C**), minus the cost of leisure (**area C**), the only social cost in **Figure 5.3**. The treatment of the opportunity cost of labour can be misleading unless the practitioner strictly follows one of the described approaches.

When the project employs workers in a segment of the labour market with minimum wage regulation, the opportunity cost of one unit of labour can be approximated, pending of more precise information, by the average of the official minimum wage (\bar{w}) and the lowest reservation wage (\underline{w}). When this last value is unknown, a lower bound for the shadow price of labour can be calculated as $(1/2)\bar{w}$, which implies a lowest reservation wage of zero.

Finally, the determination of the shadow price of labour has other complications when we look to the interactions of the impact of the project in the primary market with other markets through complementarity and substitutability relationships. The discussion above ignores the effect of the project in other sectors of the economy and hence on aggregate unemployment. The examination of other sectors closely related to the primary market is recommended.

³² Again, we are assuming that there is no economic cost of public funds, that is, the shadow price of public funds is equal to one. Otherwise, the additional benefit $\lambda_g(w^0\tau_w + u)(L^1 - L^0)$ should also be considered.

6. ECONOMIC VALUATION OF NON-MARKET GOODS

6.1. Methods for valuating external effects in transport projects

The best way to approximate the opportunity cost of the resources used in a transport project is to use the prices of the markets where those resources are exchanged. However, there are goods – such as transport safety or air quality – for which these markets are not always defined, and alternative valuation procedures must be devised for them. Even when they are not the direct objective of transport projects, the results associated with these resources often appear as side effects for other people (for example, road enhancements affect the surrounding landscape), together with other outcomes (pollution, noise, congestion, etc.) imposed on social agents not necessarily participating in the same transport markets.

Ignoring the externalities of a project on the rest of the society may conceal its true opportunity costs or, equivalently, hide the real benefit that society derives from its implementation. If the social profitability of the project is sensitive to the magnitude of the externality, it may even change the accept/reject decision or the project ranking. For these reasons, the problem addressed in this section is not only to value the resources for which market prices are not readily available, but also to understand their particular characteristics and why and how they are included in the CBA of transport projects.

This economic valuation of the costs and benefits of non-market resources associated with a transport project is usually carried out using different methods that aggregate and monetise individual preferences, using as a starting point the concepts of willingness to pay (*WTP*) and willingness to accept compensation (*WTA*), and generally assuming that, if any impact caused by a transport project is not perceived by individuals, it will not entail any change in social welfare and, therefore, will have no value from the CBA point of view.³³

However, the use of these preference-based valuation methods is not always easy and, sometimes, more direct procedures (for example, just quantifying the external impacts of a project through their restoration, avoidance or replacement costs) are used instead. The main drawback of these cost-based alternatives is that they do not consider the individuals' preferences, and implicitly assume that if those costs were actually

³³ In a context of limited information and bounded rationality, this perception includes both subjective and objective elements, and the mere existence of the externality or the choice between one way or another to measure changes in welfare basically depends on the perception by the individuals of their right to benefit from the improvement or not to bear their costs (Kahneman and Tversky, 1979). A wider (social planner) viewpoint may be advisable in these cases.

incurred it was because the society valued the resources at least that amount. Furthermore, these methods can be only applied when the costs to replace, avoid or restore the loss have been completed, and this means assuming that decision-makers made an optimal choice when deciding *ex ante* whether to undertake these expenditures.

Therefore, the most recommended techniques for assessing external effects on transport projects (those based on preferences) can generally be classified into two main groups: those that rely on transport-related markets and those using hypothetical markets. Since they have different assumptions, not all of them can be used to measure all impacts: it is necessary to consider each situation separately and clarify exactly what should be assessed when choosing one technique or another. Sometimes it may be needed to use several methods simultaneously to focus on different aspects and estimate the total impact of the externality.

6.1.1. Techniques based on transport-related markets

This set of techniques elicits the individuals' preferences from their decisions in other markets that have some kind of relationship (substitutability or complementarity) with the non-market goods. Firstly, the *averting behaviour method* is based on the idea that the *WTP* of any individual for a marginal change in the quantity or quality of a non-tradable good or service affected by a transport project can be expressed as a marginal change in expenditure on other goods or services.

Despite its conceptual attractiveness, this revealed-preference approach has several weaknesses. For example, it does not apply to non-marginal changes. According to this procedure, the benefit of a non-marginal change in any non-tradable good or service *z* (e.g. less pollution) would be valued as the reduction in expenditure on the private good *x* that makes the individual retain its level of utility (e.g. lower health care expenditure). However, since he has more resources because of the improvement in *z*, the individual may increase his consumption of all other goods and services. Thus, the expenditure reduction in good *x* could be less than the required to hold the utility level constant and therefore it will underestimate the benefits of the change in *z*. Note also that this method assumes that individuals quickly adjust their decisions to changes in *z*, when in fact some time may be needed. In addition, when preferences are complex, an averting behaviour may not fully compensate for a decline in *z* or even lead to other benefits.

The *travel cost method* is an alternative valuation technique based on analysing the complementarity relationship between a non-tradable good or service (e.g. visiting a natural landscape) and a tradable one (e.g. making a trip) which is required to consume the first one and indirectly captures the individuals' preferences for it. The idea is that the costs that people incur to visit a site can be interpreted as the 'price' for it and,

therefore, used to obtain the *WTP* for the site as a whole, or for some of its specific characteristics. Sometimes these money values may be derived from standard regression techniques (using, for example, the number of visits as explanatory variable), although they should always take into account generalised prices, the analysis of alternatives in the decision-making process, the distribution of travel costs when individuals visit more than one site, or how to aggregate visits with different durations or motives.

The *hedonic pricing method* also uses the related market approach. It is based on an alternative to the neoclassical consumer theory put forward by Lancaster (1966), according to which differentiated products can be fully described after a series of objectively measurable characteristics or attributes, so that prices reflect their differences. Again, this approach assumes complementarity between a tradable good or service and a non-tradable one, but in this case the relationship operates through price changes. In fact, the equilibrium relationship between the price of a good and its characteristics-vector is called a hedonic price function, and its partial derivatives provide the marginal implicit values of each characteristic, that is, the additional expenditure required to achieve a marginal change in any particular attribute. In a competitive market, this implicit price will equal the individual's *WTP* for the marginal change in that feature (Freeman, 1993). However, since not all individuals have the same preferences, it is also required to estimate a function that explains how the marginal *WTP* varies with the level of each attribute, the socioeconomic characteristics of individuals, and any other variables that might affect preferences.

Although it has been widely used in transport (for example, to approximate the effects of environmental externalities through changes in property values) this method faces at least two criticisms: from a theoretical viewpoint, its validity largely relies on strict assumptions about perfect information, perfect mobility of consumers, and existence of equilibrium in the transport-related market; from an empirical perspective, the estimation of hedonic price functions is critically affected by the quality of the available datasets, and biased estimates or multicollinearity problems are not infrequent.

6.1.2. Techniques based on hypothetical markets

Both as an alternative and a complement to the above techniques, there are other valuation methods that attempt to obtain consumer preferences through survey-based hypothetical transactions (markets). The corresponding estimates of changes in social welfare are not derived from the revealed behaviour of individuals but inferred from what would be their declared behaviour in social experiments, assuming that they will do exactly what they answered. To ensure that this assumption is effectively satisfied, the survey design is of great importance for all these techniques: it must be devised so

that people perceive the questions as real issues, minimising as much as possible the existence of bias and strategic behaviour on the part of the respondents.

In general, these techniques can be grouped into contingent valuation methods and multi-attribute methods. The first one focus on the holistic nature of the goods or services to be assessed, while the second group pays more attention to the attributes that define them. These differences determine the format of the survey questions and the treatment given to the collected data.

In *contingent valuation methods* individuals are offered a hypothetical choice in exchange for a certain amount of money. The type of questions varies across studies and depends on the valuation purpose. The most direct ones have ‘open format’ and directly ask individuals their *WTP* or *WTA* for the proposed change. The responses can be then analysed by estimating regression models to explain them as a function of other variables. The main disadvantage of this approach is that it places individuals in a rather unusual position: in real markets, they decide between a set of goods and services with well-defined prices, and rarely face situations where they are asked to make an offer that can be accepted or rejected by the seller. As a result, surveys using this format often provide high non-response rates and implausible extreme values.

Other question-formats often extract less information but provide more complex responses. For example, in the ‘single-bounded referendum’ format, a subsample of individuals is asked if they would make a payment for the provision of a public good, varying the amount of such payment across subsamples, which are later compared. Acceptable answers are normally restricted to closed categories that may be also analysed using regression techniques. Similarly, in a ‘double-bounded referendum’, respondents are asked whether they would pay a certain amount of money to acquire the good. If the answer is yes, then they are prompted for a higher (lower, if answered no) amount (and so on). Although this auction-like approach may simplify grouping consumers into different preferences classes, its main problem is that the responses to the subsequent questions may be influenced or be inconsistent with the previous ones. Additional (mixed) question-formats are also possible, but the choice among them often depends on each case.

Alternatively, when *using multi-attribute methods*,³⁴ individuals must express their preferences on sets of alternatives defined by attributes that vary at different levels and depend on the objective of the study. Since it is precisely the changes in the values of

³⁴ Multi-attribute decision making methods are generally restricted to a discrete and limited number of prespecified alternatives, requiring inter and intra-attribute comparisons and implicit or explicit trade-offs. They are included into a wider set of Multi-Criteria Analysis (MCA) methodologies (Zanakis *et al.*, 1998).

these attributes which are intended to be valued by these techniques, it is required that one of them represents a monetary payment. The number of attributes and levels may not be too large to minimize the difficulty of respondents in assimilating the information provided. Preliminary tests and successive pilots are often required to refine the survey and verify that the population considers the included attributes and their levels to be relevant and easy to understand. Once the attributes and the levels are selected, they are combined to obtain the different alternatives, which are randomly grouped into consistent choice sets where at least one of which represents the current situation (without the project).

Then, one or more of these sets are presented to individuals to express their preferences, either by choosing only one or by ranking a subset of them. In the first case, each respondent selects his most preferred alternative, a type of experiment that is easy to answer for individuals since it reminds them, to some extent, of the kind of tasks they must perform in real markets. However, from a researcher's point of view, this variant is the one that provides less information for each individual and set. Alternatively, in the ranking exercise, individuals must order all alternatives included in the choice set according to their preferences. This format provides more information, but it raises doubts about the ability of individuals to provide consistent answers when the number of alternatives is high, when they have similar preferences on several alternatives, or when the choice set includes alternatives that individuals would never choose or do not know fairly well. The empirical analysis of the responses to these experiments is finally based on random utility models (*logit* or *probit*), whose parameters allow us to indirectly estimate the *WTP* of individuals for marginal changes in the levels of one particular attribute.³⁵

6.2. Environmental externalities and transport projects

The incorporation of the environmental effects of a transport project into its economic assessment constitutes one of the most relevant applications of the valuation procedures for non-market goods, such as the health loss suffered by individuals due to poorer air quality caused by road transport, or the worsening of the quality of life of the residents near an airport. This process is characterized by three main assumptions: the origin of the externality lies in a production or consumption activity directly related to transport (such as building infrastructure or its utilisation for the provision of transport services), the transport activity implies a significant impact on the availability or the quality of a natural resource (land, air or water), and this impact affects the

³⁵ For non-marginal changes of one or more attributes simultaneously, a simplified approach is also available and commonly used. It requires assuming that the estimated marginal values are constant for all units and attributes, and then aggregate those values linearly. However, more formal procedures can be found in Hanemann (1982).

welfare of individuals not directly involved in the transport activity (i.e., the rest of the society).

In addition, the environmental impacts of transport can be positive for some individuals and negative for others,³⁶ they are not necessarily constant over time and, generally they are non-rival, affecting several people at the same time without varying the individual value of the effect. Thus, when valuing environmental externalities, an average value per person can be used, and then multiplied by the number of affected people. If the externality is positive for a part of the population and negative for another part, the aggregation of the individual values is assumed to reflect the overall effect on social welfare. However, externalities vary with the transport mode or infrastructure type, with each stage of the project, and with how their ultimate effects are translated to individuals. Thus, different types of external effects may require different approaches in CBA, as discussed below.

6.2.1. Noise

Noise is one of the major externalities associated with building and operating most types of transport infrastructures and services. It affects quality of life in general, work performance and recreational activities, and may even lead to deterioration in the health of some people. According to Navrud (2002), its monetary valuation usually requires the following steps:

1. Estimate a noise dispersion model relating the project incremental effects to noise exposure at different locations. These effects can be measured in decibels or any other indicator of noise levels.
2. Use dose-response functions to technically relate these noise level indicators with nuisance levels, health effects or other human impacts due to increased noise.
3. Identify different types of health impacts associated with this externality: either punctual (sudden noise) or related to prolonged exposure over time.
4. Gather information on the number of people/households affected by the project at every level of discomfort established by the dose-response function. This information will allow to finally calculate the total noise impact of the project.
5. Define money values for the units (noise levels) resulting from the dose-response function. In general, two different types of units are often used: either an average money value per decibel per person and year, or different values for each level of

³⁶ This section deals with the negative ones. Positive one (apart from the 'Mohring effect', discussed below) will be covered in **Section 7**, as agglomeration economies and other network effects.

discomfort (e.g., mild, moderate, severe). These values are usually calculated using hedonic price models (applied to the housing market) or alternative valuation methods based on stated preferences.

6. Finally, the total monetary value of the noise impact arising from the project is calculated by multiplying the total number of people potentially affected at each level of discomfort of the dose-response function by the value of each level of economic discomfort.

In the recently updated *Handbook on the external costs of transport* (European Commission, 2019), detailed noise costs associated with transport are estimated for EU-28 using a bottom-up approach. First, for each noise class and transport mode, the total number of people exposed is calculated. Then, noise costs per person (consisting of an annoyance value and a health value) are obtained using a direct *WTP* approach and an environmental burden of disease method. These values are multiplied by the exposed people, yielding the total external noise costs for each mode, which are then allocated to specific vehicle categories. Finally, average noise costs are estimated by dividing the total costs by total transport performance.

Table 6.1 summarizes the total and average noise costs for road and rail transport for passenger-km (pkm), vehicle-km (vkm) and ton-km (tkm) in 2019 euros.³⁷ **Table 6.2** illustrates the average costs of both passenger and freight aviation. Passenger aviation values are provided per landing and take-off operation (LTO), passenger and pkm. The average costs of freight aviation are provided per LTO, tonne and tkm.³⁸

³⁷ Costs at the country level are provided in a separate database by European Commission (2019). For road transport, the total costs originating from transport via a motorway are also provided in the database. In aviation, the values refer to 33 European airports. Costs at the airport level are provided in the database.

³⁸ Marginal noise costs for road and rail transport, which differ from average costs mainly because local factors that influence the noise level and the damage and annoyance level, are also provided by European Commission (2019). In aviation, they heavily depend on specific factors (such as population density around airports, flight path, aircraft type and technology, time of the day, etc.) and, hence, it is very difficult to present an accurate range of values that could be applied for all projects.

Table 6.1. Total and average noise costs for land-based transport

	TOTAL COSTS	AVERAGE COSTS	
Passenger transport	Billion €	€-cent per pkm	€-cent per vkm
Passenger car	26.2	0.6	0.9
<i>Passenger car - petrol</i>	13.8	0.5	0.8
<i>Passenger car - diesel</i>	12.4	0.6	0.9
Motorcycle	14.8	9.0	9.4
Bus	0.8	0.4	8.0
Coach	0.9	0.2	4.7
Total passenger road	42.6		
High speed passenger train	0.4	0.3	97
Passenger train electric	2.6	0.8	106
Passenger train diesel	0.9	1.4	81
Total passenger rail	3.9		
Total passenger transport	46.5		
Freight transport	Billion €	€-cent per tkm	€-cent per vkm
Light commercial vehicle	5.4	1.6	1.1
Heavy goods vehicle (3.5-7.5 t)	1.0	1.2	4.0
Heavy goods vehicle (7.5-16 t)	1.8	0.8	5.7
Heavy goods vehicle (16-32 t)	3.0	0.4	6.5
Heavy weight vehicle (> 32 t)	3.2	0.4	7.2
Total freight road	14.5		
Freight train electric	2.1	0.6	359
Freight train diesel	0.4	0.4	201
Total freight rail	2.5		
Total freight transport	17.1		
Total road, rail, inland waterway	63.6		

Source: European Commission (2019). Aggregate and average €2019 values for EU28. Average costs are given by passenger-km (pkm), vehicle-km (vkm) and ton-km (tkm). For rail passengers in electric trains the total costs do not include high-speed trains.

Table 6.2. Total and average noise costs for air transport

	TOTAL COSTS	AVERAGE COSTS			
Passenger transport	Billion €	€/LTO	€/passenger	€/tonne	€-cent/pkm
Short haul	0.84	257	2.05	9.04	0.46
Medium haul					0.11
Long haul					0.01

Source: European Commission (2019). Aggregate and average €2019 values for 33 major airports. Average costs are also provided by landing and take-off operations (LTO). No values available for freight air transport.

6.2.2. Effects on the landscape

Transport projects often lead to a degradation of the landscape or recreational value of the area in which they are developed. Since this type of impact is site-specific, it is not possible to establish comparable benchmarks. The monetary valuation of these impacts is usually performed using the *travel cost method*, although, this technique is only able to capture the value assigned to those services in the area that require people to move

there for its enjoyment, ignoring other use values. Revealed preferences methods, such as contingent valuation, may be also helpful to quantify the monetary value assigned by individuals who have not expressed their preferences by moving into the affected area.

The *Handbook on the external costs of transport* considers that the different negative effects of transport on nature and landscape can be grouped into three categories: habitat loss, associated with the land use of transport infrastructure and services when they deteriorate the ecosystems and reduce biodiversity; habitat fragmentation, when the infrastructure (e.g. a road) induce separation effects for wildlife; and habitat degradation due to emissions, related to air pollution effects (discussed below).

In general, there are limited studies covering the external costs of habitat damage due to transport activities. The total and average costs of habitat damage suggested in European Commission (2019) are thus calculated based on the European research project INFRAS, which takes into account the infrastructure network length (or area) and average cost factors for habitat loss and habitat fragmentation, as showed in **Table 6.3** for land-based modes and in **Table 6.4** for air transport.

On the other hand, according to European Commission (2019), the marginal costs of habitat loss are virtually zero (only if infrastructure capacity must be enhanced due to high demand, there are additional marginal costs). The marginal costs of habitat fragmentation, however, can be assumed to be substantial, and in some cases almost as high as the average cost of habitat fragmentation, since the traffic on a road really hinders animals to pass. However, it is not possible to make a generally applicable estimation of the marginal costs of habitat damage.

Table 6.3. Total and average costs of habitat damage for land-based transport

	TOTAL COSTS	AVERAGE COSTS	
Passenger transport	Billion €	€-cent per pkm	€-cent per vkm
Passenger car	25.9	0.55	0.9
<i>Passenger car - petrol</i>	14.1	0.54	0.9
<i>Passenger car - diesel</i>	11.8	0.56	0.9
Motorcycle	0.5	0.33	0.3
Bus	0.2	0.10	1.9
Coach	0.4	0.11	2.2
Total passenger road	27.1		
High speed passenger train	0.7	0.62	185
Passenger train electric	1.4	0.57	75
Passenger train diesel	0.5	0.84	49
Total passenger rail	2.7		
Total passenger transport	29.7		
Freight transport	Billion €	€-cent per tkm	€-cent per vkm
Light commercial vehicles	4.4	1.35	0.9
Heavy goods vehicles	3.6	0.19	2.4
Total freight road	8.0		
Freight train electric	0.8	0.24	134
Freight train diesel	0.2	0.25	111
Total freight rail	1.0		
Inland Vessel	0.3	0.20	2.9
Total freight transport	9.3		
Total road, rail, inland waterway	39.1		

Source: European Commission (2019). Aggregate and average €2019 values for EU28. Average costs are given by passenger-km (pkm), vehicle-km (vkm) and ton-km (tkm). For rail passengers in electric trains the total costs do not include high-speed trains.

Table 6.4. Total and average costs of habitat damage from air transport

	TOTAL COSTS	AVERAGE COSTS	
Passenger transport	Billion €	€-cent/pkm	€-cent/pax
Short haul (< 1,500 km)	0.024	0.027	14.9
Medium haul (1,500-5,000 km)	0.021	0.007	12.3
Long haul (> 5,000 km)	0.005	0.0008	6.5
Total passenger aviation	0.050	0.007	12.2
Freight transport	Billion €	€-cent/tkm	
Total freight aviation	0.006	n.a.	

Source: European Commission (2019). Aggregate and average €2019 values for 33 major airports. Average costs are also provided by landing and take-off operations (LTO).

6.2.3. Air pollution

In addition to its direct effects on the quality of life and human health, the increase in air pollution associated with transport projects may cause other impacts on production and consumption activities, such as reduced visibility or deterioration of buildings or other exposed materials. Its measurement and treatment in CBA share some common features with noise effects and may be summarized in the following steps:

1. Identification of the changes in emissions levels due to the project in all its stages. This requires detailed technical information on the project characteristics, transport flows, and the variability of emissions.
2. Calculation of the dispersion and concentration of emissions to analyse the behaviour of pollutants. This is technically complex because the relationship between emissions and pollutant concentration is non-linear and depends on many local parameters (meteorology, topography, initial level of concentration, etc.).
3. Measurement of the emissions impacts on different receptors, according to suitable dose-response functions, particularly focusing on human health (mortality and morbidity), material damage (infrastructure) and verifiable losses in agricultural production and forestry.
4. Monetary valuation of such impacts, according to their characteristics. For example, for the assessment of crop losses and material damage, market prices can be used. However, the monetary valuation of health changes is more difficult, since it should include estimates on the medical expenses to be incurred, opportunity costs from lost productivity of the affected people, and any other disutility caused by the impact of air pollution on health. Some of these elements can be monetized using market prices, but the last component, usually the most relevant, requires estimating individuals' *WTA* to prevent the occurrence of damage to their health or their *WTP* to reduce the risk of dying due to emissions.

This four-step procedure suggests that precise estimates for air pollution costs are not always feasible, given the amount of resources and time that would otherwise be needed. Therefore, it is usual to simplify the valuation process (by reducing the number of pollutants or by simplifying their modelling) and to adapt estimates from other studies. The main disadvantage of these extrapolations is that the pollution impacts caused by transport projects are largely project specific. Despite this, as suggested in the *Handbook on the external costs of transport*, the existence of a broad range of international studies (particularly in Europe) provides some references for total and average air pollution costs for all transport modes for the EU28, as showed in **Table 6.5**.

Table 6.5. Total and average air pollution costs (all transport modes)

	TOTAL COSTS	AVERAGE COSTS	
Passenger transport	Billion €	€-cent/pkm	€-cent/vkm
Passenger car	33.36	0.71	1.14
<i>Passenger car - petrol</i>	8.58	0.33	0.53
<i>Passenger car - diesel</i>	24.79	1.18	1.90
Motorcycle	1.84	1.12	1.17
Bus	1.35	0.76	14.19
Coach	2.67	0.73	14.34
Total passenger road	39.23		
High speed passenger train	0.002	0.002	0.66
Passenger train electric	0.03	0.01	1.14
Passenger train diesel	0.52	0.80	47.0
Total passenger rail	0.55		
Total passenger transport	39.78		
Freight transport	Billion €	€-cent/tkm	€-cent/vkm
Light commercial vehicle	15.49	4.68	3.24
<i>Light com. vehicle (petrol)</i>	0.33	1.72	1.17
<i>Light com. vehicle (diesel)</i>	15.16	4.86	3.37
Heavy goods vehicle	13.93	0.76	9.38
Total freight road	29.42		
Freight train electric	0.01	0.004	2.14
Freight train diesel	0.66	0.68	305.39
Total freight rail	0.67		
Inland Vessel	1.93	1.29	1,869
Total freight transport	32.02		
Total road, rail, inland waterway	71.80		
Air transport	Billion €	€-cent/pkm	€-cent/pax
Short haul (< 1,500 km)	0.27	0.30	163
Medium haul (1,500-5,000 km)	0.38	0.13	231
Long haul (> 5,000 km)	0.36	0.06	444
Total air transport	1.01	0.10	246
Maritime transport	Billion €	€-cent/tkm	
Freight shipping	29	0.4	
Total maritime transport	29		

Source: European Commission (2019). Aggregate and average €2019 values for EU28. Average costs are given by passenger-km (pkm), vehicle-km (vkm) and ton-km (tkm). Air transport costs are calculated for 33 selected airports. Maritime data are only available for freight shipping.

6.2.4. Climate change

The impact of transport on climate change is mainly associated with the emission of greenhouse gases (GHG) such as carbon dioxide (CO₂), methane (CH₄) or nitrous oxide (N₂O). However, its assessment is highly complex because it implies long-term effects of a very different nature (floods, forest fires, heat or cold waves, droughts and other natural disasters) and great uncertainty about their specific location and cause-effect relationship.

The usual approach to the valuation of climate change effects in CBA requires first to quantify the increase in emissions of the different greenhouse gases caused by the project. Given the global character of this impact, it is not relevant to determine where and how the emissions occur. Instead, this total amount is then converted into 'equivalent emissions of CO₂', used as a standard reference, and finally this value is multiplied by a benchmark external cost per tonne of CO₂ to get an estimate of the overall externality.

There are two main specific methodologies to obtain this unit external costs associated with climate change (Maibach *et al.*, 2008). The first one is based on the costs of damage, which tries to establish the physical impacts of climate change and then proceeds to the monetary valuation of such damage. The main disadvantages of this method lie in the difficulty of identifying the impacts to be assessed, in predicting their long-term developments, in distributing different impacts across different regions and in establishing the time horizon for the analysis, and the adequate social discount rate. An alternative approach is based on the costs that would be required to achieve a given reduction in GHG emissions. The challenges posed by this procedure are associated with setting the targets and determining the concerned sectors (or countries). It has also the disadvantage that the monetary values are not related to the individuals' preferences of individuals.

In European Commission (2019), total and average climate costs are calculated using a bottom-up approach for road, rail, inland waterway, aviation and maritime transport, where the initial inputs are estimated GHG emission factors per vehicle type, vehicle performance data and the climate change costs per tonne of CO₂ equivalent. **Table 6.6** summarizes these costs for all transport modes and EU28, although marginal costs values are also provided in the *Handbook* for selected cases.³⁹

³⁹ Notice that including both CO₂ taxes and climate costs means a double-counting. In the case of air transport, most flights within the EEA require emission permits (EU Emission Trading System). To include both the permit cost and climate costs also means a double-counting.

Table 6.6. Total and average climate change costs (all transport modes)

	TOTAL COSTS	AVERAGE COSTS	
Passenger transport	Billion €	€-cent per pkm	€-cent per vkm
Passenger car	55.56	1.18	1.90
<i>Passenger car - petrol</i>	32.02	1.22	1.97
<i>Passenger car - diesel</i>	23.54	1.12	1.80
Motorcycle	1.47	0.89	0.94
Bus	0.84	0.47	8.83
Coach	1.61	0.44	8.66
Total passenger road	59.49		
Passenger train diesel	0.22	0.34	20.1
Total passenger transport	59.71		
Freight transport	Billion €	€-cent per tkm	€-cent per vkm
LCV	13.17	3.98	2.75
<i>LCV - petrol</i>	0.71	3.76	2.56
<i>LCV - diesel</i>	12.45	3.99	2.77
HGV	9.63	0.53	6.48
Total freight road	22.79		
Freight train diesel	0.24	0.25	112.4
Inland Vessel	0.40	0.27	383.1
Total freight transport	23.43		
Total road, rail, inland waterway	83.14		
Air transport	Billion €	€-cent/pkm	€-cent/pax
Short haul (< 1,500 km)	2.14	2.39	1,315
Medium haul (1,500-5,000 km)	5.50	1.85	3,341
Long haul (> 5,000 km)	14.37	2.24	17,629
Total air transport	22.01	2.14	5,383
Maritime transport	Billion €	€-cent/tkm	
Freight shipping	11	0.16	
Total maritime transport	11		

Source: European Commission (2019). Aggregate and average €2019 values for EU28. Average costs are given by passenger-km (pkm), vehicle-km (vkm) and ton-km (tkm). Air transport costs are calculated for 33 selected airports. Maritime data are only available for freight shipping.

6.2.5. Other environmental costs of transport

There are other environmental costs of transport whose measurement is not always simple nor direct. For that reason, these costs are often addressed using a qualitative approach, describing their main characteristics in each case and suggesting general procedures for their inclusion in the CBA. For example, the environmental impact of soil contamination may have effects on plants, agricultural productivity and even on human and animal health in the long term, which represents a source of additional complexity.

Sometimes this is avoided by using cost-based techniques (valuating the impact according to the cost to replace or restore the environmental good) but this ignores the individuals' preferences and may result in arbitrary values not related to the real social costs. If possible, a more appropriate approach is to try to quantify particular effects (on selected people or areas) by using dose-response functions (as in the case of noise or pollution) for their subsequent valuation using market prices (for example, in the case of effects on soil productivity) or stated or revealed preferences methods (for the effects on human health). Similar procedures are valid for water contamination, whose effects (including damage to wildlife and to commercial and recreational activities) may appear long afterwards, and not around the initial point of contamination.

Finally, with respect to vibrations, it is a very specific, location-based impact that is often ignored in many transport projects. However, the inconveniences that they cause to human receptors within its sphere of influence may be approached as in the case of noise; their negative impact on infrastructure or on specific production activities could be valued using market prices.

6.3. Accident costs and the statistical value of human life

Either by mechanical failures, external conditions or, more commonly, by the influence of human errors, accidents are relatively common events in transport activities. They usually result in substantial costs including at least two types of components: material costs (e.g. damage to vehicles, administrative fees and medical costs) and immaterial costs (e.g. shorter lifetimes, suffering, pain and sorrow). Market prices can be used to evaluate the first component, but no such prices exist for immaterial costs. In addition, part of the total costs of accidents is already internalised, for example through insurance premia or through accounting for risks that are well anticipated. These features imply that accidents share some of the elements associated with the assessment of environmental effects, but also have their own distinctive characteristics.

From the point of view of CBA, a typical benefit of many transport projects is the reduction of the risk arising from the use of certain infrastructures or services. Their

proper maintenance or improvement often make trips safer, ultimately resulting into fewer fatalities or less serious injuries. These changes imply an improvement in social welfare that must be incorporated in the evaluation, regardless of the benefit derived from the cost savings that avoiding an accident could also cause. These effects are commonly evaluated by analysing the individuals' trade-offs between marginal changes in their risk levels and their income, either in related markets (usually, the labour market, where hedonic wages models can be applied to study whether an individual would accept a riskier job in exchange for a higher wage) or in hypothetical markets (using surveys to reveal preferences about risk).

Although there is no standard concept for the value of a specific human life in economics, the measurement of the immaterial costs associated with changes in the risk of accidents is usually assessed using the 'value of a statistical life' (*VSL*), which may differ across modes, accident circumstances and countries.⁴⁰ In the US and Canada, for example, this *VSL* is estimated using labour market studies, with values ranging between €6.9 million and €4.1 million, respectively. In Europe, *VSL* is generally approached using *WTP* studies, typically in terms of the 'value of a life year' (*VOLY*), defined as the amount of money that people are willing to pay for one year of additional life expectancy. The usual calculation procedure requires confronting the individuals with successive monetary valuations which are increased in a proportion equivalent to that required to transform small changes in risk levels associated with a certain probability of death (probability equal to one). For example, if individuals were willing, on average, to reduce by 1 per 10,000 their risk of dying in an accident in exchange for €100 every year, this would imply a value of a statistical life of €1 million. The main problem with this approach is that it produces estimates of the monetary value of changes from 0 to 1 on the probability of fatal accidents by using estimates of monetary values obtained from marginal changes in risk.

On average, the resulting *VSL* in Europe is €3.6 million (in 2016 prices) but, taking into account income differences across countries and deducing consumption loss (to avoid double counting with gross production loss), the human costs of fatalities for the EU28 recommended in European Commission (2019) is €2.9 million. The human costs of injuries are valued at 13% and 1% of the *VSL* respectively for serious and slight injuries.

⁴⁰ A recent review of technical procedures and *VSL* values across the world can be found in Viscusi (2018).

Table 6.7 summarizes the total and average external accident costs in the EU28 for all transport modes. Disaggregate costs at the country level and for selected airports and ports are provided in the accompanying database.⁴¹

Table 6.7. Total and average accidents costs for all transport modes

	TOTAL COSTS	AVERAGE COSTS			
Passenger transport	Billion €	€-cent per pkm			€-cent per vkm
Passenger car	210.2	4.5			7.2
Motorcycle	21.0	12.7			13.3
Bus/Coach	5.3	1.0			18.9
Total passenger road	236.5				
High speed passenger train	0.1	0.1			17.3
Conventional passenger train	2.0	0.5			52.2
Total passenger rail	2.0				
Total passenger transport	238.5				
Freight transport	Billion €	€-cent per tkm			€-cent per vkm
Light commercial vehicle	19.8	6.0			4.1
Heavy goods vehicle	23.0	1.3			15.5
Total freight road	42.8				
Freight train	0.3	0.1			34.1
Inland Vessel	0.1	0.1			86.3
Total freight transport	43.1				
Total road, rail, inland waterway	281.7				
Air transport	Million €	€/LTO	€/pass	€/tonne	€-cent/pkm
Short haul	75.01	22.95	0.18	0.81	0.04
Medium haul					0.01
Long haul					0.001
Maritime transport	Million €	€ per port call			€ per mill. pass.
Passenger ship	3.3	26			40,996
Freight maritime transport	Million €	€ per port call			€ per million tonnes
Freight ship	63.3	318			36,524

Source: European Commission (2019). Aggregate and average €2016 values for EU28. Average costs are given by passenger-km (pkm), vehicle-km (vkm) and ton-km (tkm). Air transport costs are calculated for 33 selected airports with data per landing and take-off operations (LTO). Costs per passenger are including the complete flight (not only the half-way principle). Maritime data are only available for freight shipping.

⁴¹ Marginal accident costs are only provided in the *Handbook* for road transport. For all other modes of transport, the marginal accident costs are considered to be equal to the average costs, assuming that the other modes are scheduled services, which implies that the accident risk is less dependent on the amount of traffic for these modes.

6.4. Congestion as an externality

Congestion is a condition that appears in some transport modes as the number of users of a given infrastructure increases and vehicles are progressively delayed when travelling. Slower average speeds and longer queues increase users' travel time and their generalised price, reducing social welfare. It is typically associated with urban and interurban road transport, where congestion costs can be readily calculated in terms of speed-flow ratios. This approach cannot be extended to scheduled transport modes, as they essentially provide services which are already planned according to the allocative capacity of networks and nodes. However, in the case of bus or train stations, or port or airport terminals, there may also be congestion in some critical processes for handling passengers or cargo (security controls, customs, boarding, etc.).

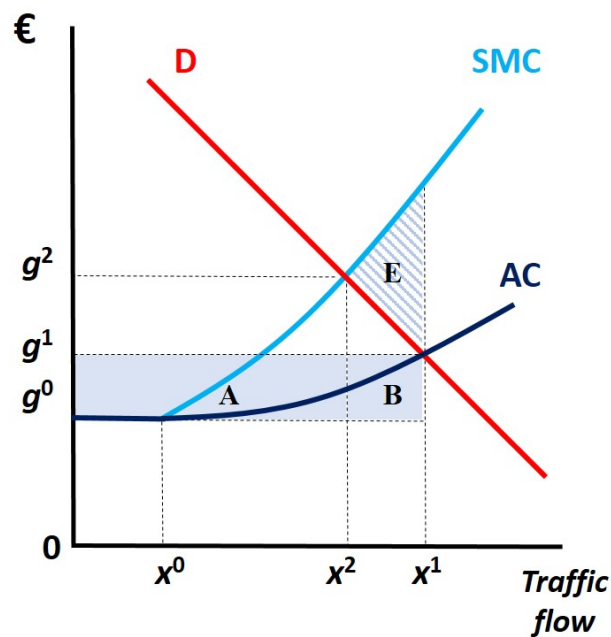
Interestingly, it has also been observed (Mohring, 1972) that there may exist an opposite positive externality related to reduced travel time when the frequency (e.g. buses per hour) of scheduled services increases with the number of users (due to shorter waiting times). Although its empirical validation is not always easy, this 'Mohring effect' is often argued in support of public transport subsidies, on the grounds that they are required to achieve marginal cost pricing, since the average cost of a passenger-journey includes the average waiting time, while his (lower) marginal cost includes only the average waiting time less the diminution in total waiting time caused by the increased frequency.

In any case, congestion can be considered as an externality insofar as it is an effect generated by some users who do not take into account the costs that they impose on the rest of the infrastructure users. However, it also has 'internal' effects, as all users affected by congestion are cause and effect of the same situation (although not all suffer from it equally). Because of this feature, congestion costs are sometimes excluded from an explicit account of transport externalities in CBA, and they are simply accounted for in the changes of travel time. While this scheme may be appropriate when drawing up social accounts for the transport sector, the magnitude of these costs means that the effects of congestion cannot be overlooked in project assessment. In addition, the search for solutions to reduce congestion costs is often one of the main concerns of transport policy makers.

The approach used in European Commission (2019) focuses mainly on road congestion costs, which are estimated using two different methods, the *delay cost* and the *deadweight loss (DWL)*. The speed-flow relationship is explicitly modelled as in **Figure 6.1**, where the horizontal axis represents vehicles per hour and the vertical axis reflects users' costs, in terms of their generalised price, which includes travel time and the value of time (assumed constant across users). A road with a certain maximum capacity per hour is represented. From left to right, if the traffic flow is lower than the

capacity, vehicles travel at free-flow speed, and users' cost is equal to g^0 . As the flow increases, travel time increases as a result of congestion (after $x > x^0$) and the average travel cost borne by users increases according to the shape of their (private) cost function, AC, until it intersects the demand curve, D. The 'delay cost approach' defines the road congestion cost as the value of the travel time lost relative to a free-flow situation, that is, the rectangles **A+B**.

Figure 6.1. Road congestion costs: delay costs and deadweight loss



Source: Adapted from European Commission (2019).

On the other hand, SMC represents the social marginal cost function, which is equal to the average travel cost borne by the road users AC plus the cost of the additional travel time, generated by the marginal vehicle that reduces the speed of all the other vehicles. The 'deadweight loss approach' departs from the socially optimal solution (where the demand function and the social marginal cost function intersect), and considers that the external cost of congestion is given by the costs associated with the demand in excess with respect to x^2 (**area E**).

Costs estimates using this deadweight loss are regarded as a proper basis for congestion pricing, whereas, the delay cost approach is generally more useful from the point of view of CBA, as it reflects the total congestion cost in a way which is (partly) comparable to the total costs for other externalities. In both cases, the average

congestion costs can be obtained dividing the total cost estimates by the corresponding urban and interurban traffic flows (in vkm).

For scheduled transport modes, the probability of ‘increased travel time’ is related to existing scheduling and slot assignment procedures. Thus, the external cost in these cases can be further classified according to two categories (European Commission (2019)):

- ‘congestion costs’, when one scheduled service delays another. Although the timetables will be designed to prevent this from happening, it could be the case that at high levels of utilisation, the presence of an additional scheduled service may lead to an additional delay to others, and
- ‘scarcity cost’, when the presence of a scheduled service prevents another scheduled service from operating or requires it to take an inferior slot. Therefore, scarcity costs are incurred whenever a slot is reserved. Scarcity costs denote the opportunity cost to service providers for the non-availability of desired departure or arrival times.

The European *Handbook on External Transport Costs* explicitly states that estimating all these congestion costs requires a great deal of information (e.g. traffic density, schedules combination, reliability indices, average delays, etc.) as well as complex preparations and adjustments. The necessary estimates are very context-specific and therefore very sensitive to different traffic situations. For these reasons, benchmarks are only provided for the total yearly road congestion costs for EU28 (plus Norway and Switzerland), using both the delay cost and deadweight loss approaches, adjusted according road types, different capacity levels and other simplifications, as showed in **Table 6.8** and **Table 6.9** for all general roads and interurban motorways.

Table 6.8. Total and average road congestion costs (all roads)

	DELAY COST APPROACH			DEADWEIGHT LOSS APPROACH		
	TOTAL COSTS	AVERAGE COSTS		TOTAL COSTS	AVERAGE COSTS	
Passenger transport	Billion €	€-cent/pkm	€-cent/vkm	Billion €	€-cent/pkm	€-cent/vkm
Passenger car	196.1	4.2	6.7	33.5	0.7	1.1
<i>Urban</i>	160.8	11.0	17.7	28.0	1.9	3.1
<i>Inter-urban</i>	35.3	1.1	1.7	5.5	0.2	0.3
Bus/ Coach	4.5	0.8	15.9	0.8	0.1	2.7
<i>Urban</i>	3.9	1.8	35.5	0.7	0.3	6.1
<i>Inter-urban</i>	0.5	0.2	3.1	0.1	0.0	0.5
Total passenger	200.6			34.3		
Freight transport	Billion €	€-cent/tkm	€-cent/vkm	Billion €	€-cent/tkm	€-cent/vkm
Light commercial vehicle	55.5	16.8	11.6	9.4	2.8	2.0
<i>Urban</i>	46.5	39.6	27.4	8.0	6.8	4.7
<i>Inter-urban</i>	9.0	4.2	2.9	1.4	0.7	0.5
Heavy Goods Vehicle	14.6	0.8	10.9	2.5	0.1	1.8
<i>Urban</i>	11.6	2.5	34.1	2.0	0.4	6.0
<i>Inter-urban</i>	3.0	0.2	3.0	0.5	0.0	0.5
Total freight	70.1			11.9		
Total road transport	270.7			46.2		

Source: European Commission (2019). Aggregate and average €2019 values for EU28 are calculated using both the delay cost and the deadweight loss approaches. Average costs are given by passenger-km (pkm), vehicle-km (vkm) and ton-km (tkm).

Table 6.9. Total and average road congestion costs (inter-urban motorways)

	DELAY COST APPROACH			DEADWEIGHT LOSS APPROACH		
	Billion €	€-cent/pkm	€-cent/vkm	Billion €	€-cent/pkm	€-cent/vkm
Passenger transport	Billion €	€-cent/pkm	€-cent/vkm	Billion €	€-cent/pkm	€-cent/vkm
<i>Passenger car - inter-urban</i>	3.2	0.28	0.45	0.7	0.06	0.10
<i>Coach - inter-urban</i>	0.1	0.05	0.89	0.02	0.01	0.20
Total passenger	3.3			0.7		
Freight transport	Billion €	€-cent/tkm	€-cent/vkm	Billion €	€-cent/tkm	€-cent/vkm
<i>Light vehicles - inter-urban</i>	0.9	1.2	0.83	0.2	0.3	0.19
<i>Heavy vehicles - inter-urban</i>	0.4	0.06	0.88	0.1	0.01	0.20
Total freight	1.3			0.3		
Total motorways transport	4.6			1.0		

Source: European Commission (2019). Aggregate and average €2019 values for EU28 are calculated using both the delay cost and the deadweight loss approaches. Average costs are given by passenger-km (pkm), vehicle-km (vkm) and ton-km (tkm).

Regarding scheduled transport modes, the *Handbook* does not identify many useful sources for comparable congestion or scarcity costs estimates. In fact, it is argued that they should be mostly internalised in service planning and infrastructure pricing (see **Section 4**).⁴² In railways, for example, well-established methodologies exist at country level to price tracks capacity according to service characteristics. However, there is no straightforward evidence that the actual charges always reflect the scarcity of slots. The congestion costs of a rail network can be also estimated starting from the information on the actual reactionary delays of trains, multiplied by the number of affected passengers and by a suitable average value of time. Excluding delay compensations paid to users, there are not many costs borne by infrastructure managers and operators because of congestion. However, including these compensations and other congestion costs estimates would represent double-counting.

Although congestion in urban public transport services generally receives less consideration, it may be an important source of disutility on peak times, or when projects shift users from private transport, increasing the crowding of buses, trams or the subway network and reducing social welfare for all other passengers. This externality may also lead to a suboptimal usage of the public transport system and used to justify congestion taxes. Again, the main problem with this issue is that it is very city-specific, and comparable values are difficult to obtain.

In air transport, congestion costs are associated with the lack of enough capacity to accommodate the required demand. Although aircraft movements are scheduled in advance using time slots, any perturbation introduced by exogenous factors (bad weather, strikes...) causes cascade-effects and accumulation of delays. Congestion costs by airline and airport can be estimated using the information on actual (or average) flight delays of flights, translating then into monetary terms using their corresponding direct costs (e.g. additional fuel during extended taxiing and circling periods, extra time paid to crew and ground staff, etc.) and indirect ones (lost revenues). For passengers, if the delay time per flight is available, it has to be multiplied by the number of passengers affected, using actual figures or approximating them by the aircraft capacity and average load factors.

Finally, with regard inland waterway and maritime transport are concerned, no illustrative quantification of marginal congestion costs could be identified in European Union (2019). In fact, Christidis and Brons (2016) indicate that congestion costs of

⁴² Since the estimation of scarcity costs is very complex, it has been suggested (Quinet, 2003) the use of auctions to reveal users' *WTP* for capacity, a procedure already applied by some airlines to deal with overbooking problems.

freight transport for both inland waterways and short sea shipping can be assumed to be negligible for most EU Member States.

6.5. The value of time savings

The relevance of the value of time (VoT), not only for estimating congestions costs, but also as key parameter for assessing changes in generalised prices, make it important to discuss some of its particular features in the context of CBA. In general, the value of these time (savings) can be approximated using the relevant hourly wage for working time (see **Section 2**), or a proportion of household income per capita for non-working time. Alternatively, time savings can be also valued using contingent valuation methods to estimate travel users' *WTP* for faster travel. While theoretically rigorous, this second and more specific approach is only recommended for projects where time savings constitute the main benefit and more precise (and costly) measurement are required.

There are additional considerations that may be useful in the practical estimation of the value of time. Firstly, due to productivity changes, it may not remain constant over time; instead, it is often assumed that it grows at the same rate as real per capita income, but a sensitivity analysis on this assumption may be advisable. It is also possible to consider different values depending on travel time components. In general, it is often recommended to use 'waiting time' values and 'access time' values above the 'in-the-vehicle' value of time (typically increased by factors around 2, or even more). Similarly, to account for congestion, a 50% premium over the normal value of time is a rule of thumb recommended for the European Union on the grounds that time is more highly in traffic jams or overcrowded buses conditions, for example. Other references also suggest explicitly taking into account other factors such as comfort, convenience, reliability or safety, although consensus on this issue is weak.⁴³ The following tables finally summarize these recommendations for Spain and the EU, using HEATCO (2006) as the main and most comprehensive source of reference.

⁴³ Nowadays, passengers can work, watch movies, etc. while travelling. This feature will be even more pronounced in the future if we consider the possibility of travelling in autonomous vehicles. All these facts could considerably affect the value of travel time.

Table 6.10. Travel time savings values in working time (passengers)

Air transport		Bus		Car/Train	
Spain	UE average	Spain	UE average	Spain	UE average
30.77	32.80	17.93	19.11	22.34	23.82
(35.74)	(32.80)	(20.83)	(19.11)	(25.95)	(23.82)

Source: Adapted from HEATCO (2006). Values are expressed in 2002 euros per passenger-hour. The terms in parenthesis are PPP-adjusted. UE average does not include Romania, Bulgaria and Croatia.

Table 6.11. Travel time savings values for non-working time (passengers)

Commuter Short distance					
Air		Bus		Car/Train	
Spain	UE average	Spain	UE average	Spain	UE average
12.72	12.65	6.12	6.10	8.52	8.48
(14.77)	(12.65)	(7.11)	(6.10)	(9.90)	(8.48)
Other Short distance					
Air		Bus		Car/Train	
Spain	UE average	Spain	UE average	Spain	UE average
10.66	10.61	5.13	5.11	7.15	7.11
(12.38)	(10.61)	(5.96)	(5.11)	(8.30)	(7.11)
Commuter Long distance					
Air		Bus		Car/Train	
Spain	UE average	Spain	UE average	Spain	UE average
16.33	16.25	7.87	7.83	10.94	10.89
(18.96)	(16.25)	(9.14)	(7.83)	(12.71)	(10.89)
Other Long distance					
Air		Bus		Car/Train	
Spain	UE average	Spain	UE average	Spain	UE average
13.69	13.62	6.59	6.56	9.18	9.13
(15.90)	(13.62)	(7.66)	(6.56)	(10.66)	(9.13)

Source: Adapted from HEATCO (2006). Values are expressed in 2002 euros per passenger-hour. The terms in parenthesis are PPP-adjusted. UE average does not include Romania, Bulgaria and Croatia.

Table 6.12. Travel time savings values for freight transport

Air transport		Road		Railways	
Spain	UE average	Spain	UE average	Spain	UE average
n.a.	n.a.	2.84	2.98	1.17	1.22
n.a.	n.a.	(3.30)	(2.98)	(1.36)	(1.22)

Source: Adapted from HEATCO (2006). Values are expressed in 2002 euros per ton-hour. The terms in parenthesis are PPP-adjusted. UE average does not include Romania, Bulgaria and Croatia.

7. INDIRECT EFFECTS AND WIDER ECONOMIC BENEFITS OF TRANSPORT PROJECTS

7.1. Direct and indirect benefits in transport cost-benefit analysis

The consumer surplus based calculation of conventional CBA capture potential benefits of transport schemes that are generated for both new and existing users of the transport system. These can arise via changes in the generalised price of travel, that is in time and fare / operating costs or in the quality of transport services. These are the so-called Direct User Benefits (DUBs) of a scheme and they typically constitute the largest component of benefits within conventional CBA calculations.

From theory, we know that under conditions of perfect competition, constant returns to scale, and in the absence of market failures; DUBs will capture all economic impacts of a transport improvement. In practice, however, market failures and scale economies tend to be prevalent in the spatial economy and this has led to developments in CBA methodology to capture what are referred to as Wider Economic Benefits (WEBs) (e.g. Venables, 2007, Mackie *et al.*, 2012, Graham and Gibbons, 2019). WEBs are benefits that are additional to conventional user benefits precisely because they arise from sources of market failure and because they go beyond the so-called indirect benefits.

The indirect benefits, as well as the WEBs, needs some market distortions to play some role in the economic evaluation of projects. The effects of transport improvements in secondary markets are irrelevant (in efficiency terms) if the rest of the economic is perfectly competitive. The indirect effects in transport projects goes from secondary non-transport markets to intermodal effects. In both cases, when time savings in the primary market is an intermodal effect. In both cases, the products of the secondary markets are complements or substitutes of the primary transport market. The treatment of these so called 'indirect effects' are similar for any secondary market (Harberger, 1965, Mohring, 1976).

The common practice is to ignore the indirect effects, under the assumption of perfectly competitive markets or the reality of different effects of different sing along the economy and the similar second order general equilibrium effect of alternative investments. Intermodal effects could be treated as a unique primary transport market, or if considered as indirect effects, included in the economic evaluation given their relevance in the final profitability of the project. The secondary intermodal effects can be positive or negative depending on the sign of the distortion and the cross elasticity but in the case of optimal pricing, like road congestion pricing or airport congestion charges, optimally designed, there are no additional benefits in these markets (de Rus, 2011).

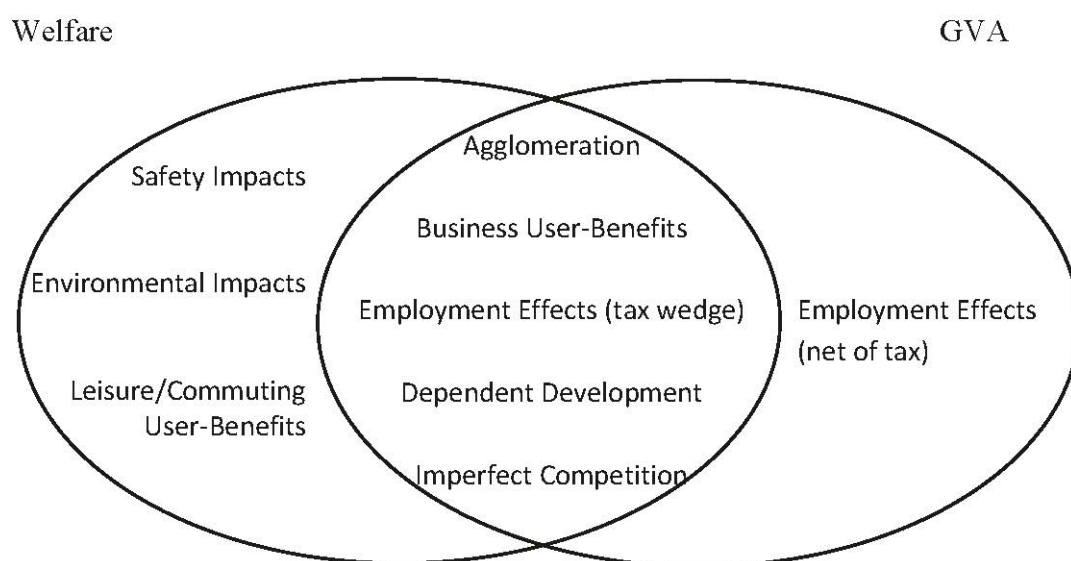
7.2. Source of Wider Economic Benefits

In current CBA practice three categories of WEBs are typically recognised.

1. **Imperfect competition** - transport improvements can cause a decrease in the costs of interacting in the spatial economy, thus potentially allowing firms to profitably expand output. Output expansion yields a welfare gain in imperfectly competitive markets when willingness to pay for the increased output exceeds the cost of producing it.
2. **Tax revenues arising from labour market impacts** - the decisions that firms and workers make about where to locate is influenced by the accessibility offered through transport systems. If accessibility improves and causes firms / workers to move to more productive locations, or otherwise have greater participation in labour markets, this will result in a tangible financial gain (i.e. higher wages or productivity). Most of this gain is captured in the consumer surplus based calculations of user benefits, but not the resulting change in tax revenue to the government (i.e. income tax, national insurance, and corporation tax).
3. **Agglomeration economies** - transport improvements can increase the potential scale of economic interactions available in the economy, with implications for the relative level of agglomeration experienced by firms. Essentially, improved transportation increases accessibility to economic mass and this yields scale economies of agglomeration.

These three sources of benefits of course have impacts on net Gross Domestic Product (GDP), or on Gross Value Added (GVA), and in fact as shown in the figure below their impacts on welfare are equivalent to those on GDP.

Figure 7.1: The links between welfare and Gross Domestic Product



Source: DfT (2016).

Figure 7.1 illustrates that GDP and welfare are not additive. The only benefits that represent additions to welfare over and above DUBs arise from distortions and failures in secondary markets. These include effects via imperfect competition, tax wage from increased employment (labour market effect), and agglomeration.

7.3. Incorporating Wider Economic Benefits in transport appraisal

Of the three sources of WEBs define above, most attention in current appraisal practice has focused on productivity effects that arise via agglomeration economies. In the UK, WEBs from imperfect competition are not typically assessed and those from labour market effects are valued via a fixed uplift that is not scheme specific (see DfT, 2016).

The concentration on agglomeration effects has arisen because these are thought to be by far the largest source of WEBs and because they can be quantified with a reasonable degree of accuracy via established econometric methods (for a review see Graham and Gibbons, 2019).

Below we explain why agglomeration or access to economic mass, is relevant for economic productivity and how transportation can affect this basic relationship.

7.4. The link between agglomeration, productivity and transport provision

A ubiquitous feature of the distribution of economic activity is a tendency towards spatial concentration, or agglomeration. Economic theory states that the process of agglomeration is driven by the presence of tangible productive advantages in the form of urban agglomeration economies. These external benefits include improved opportunities for labour market pooling, knowledge interactions, specialisation, and the sharing of inputs and outputs (e.g. Duranton and Puga, 2004). These 'mechanisms' or 'sources' of agglomeration economies are thought to result in higher productivity and lower average costs for firms. Furthermore, due to increasing returns, the larger the scale of agglomeration the greater the productivity benefits that accrue.

There is an inherent relationship between transport and the externalities of agglomeration. The prevalence of transportation costs is crucial in generating tendencies towards spatial concentration, and in fact, the level of access to economic mass experienced by any firm is largely dependent on the nature of transport provision. This is because transport systems to some extent determine proximity, or the ease of access, to other firms and to labour markets. In effect, transport can change the level of agglomeration experienced by firms and workers by rendering a larger scale of activity more accessible to them.

From this line of reasoning it is clear that there may be consequences of transport investment that relate specifically to agglomeration. Essentially, the argument is that if there are increasing returns to spatial concentration, and if transport in part determines the level of concentration or density experienced by firms and workers, then investment in transport may induce some shift in productivity via the externalities of agglomeration.

7.5. Valuing Wider Economic Impacts of agglomeration in transport appraisal

To value transport induced productivity effects from change in agglomeration we need to calculate quantity changes and value changes.

- i. **Quantity change** – the change in agglomeration caused by a transport improvement
- ii. **Value change** – the response of productivity to a change in agglomeration (i.e. agglomeration elasticity).

Below we discuss each of these in turn.

7.5.1. Measuring change in agglomeration

Measures of agglomeration are typically computed as metrics representing Access to Economic mass (ATEM). Commonly, this is done in the form of an effective density (*ED*) measure of ATEM at the zone level, where from any zone *i* the *ED* is given by

$$ED_i = \sum_{j=1}^n \frac{M_j}{d_{ij}^\alpha},$$

where *M* is some measure of economic mass at zone *j* such as employment, *d_{ij}* is a measure of the impedance between zone *i* and *j* (e.g. distance, travel time, or generalised cost), and *α* is a parameter that determines the spatial decay of agglomeration (e.g. Graham and Gibbons, 2019).

Transport improvements can change the *ED* metric via two routes.

- i. **Static change in agglomeration** – by improving the transport system travel times, and thus generalised costs (GCs) will fall leading to an increase in agglomeration via a reduction in impedance. This so called ‘static’ change in agglomeration works only through impedance, the economic mass at each zone is assumed constant before and after the transport improvement.
- ii. **Dynamic change in agglomeration** – by improving the transport system economic activities may choose to switch location, thus altering the masses at zones. ‘Dynamic’ changes encapsulate change in the physical distribution of agglomeration (or the spatial land use equilibrium).

To assess static changes in agglomeration a transport model is used to predict changes in travel times or GC, and thus changes in impedance. To assess dynamic changes a land use, or land use transport interaction model (LUTI) model, may be required to predict how the land use equilibrium will look following the transport improvement.

7.5.2. Estimating the elasticity of productivity with respect to agglomeration

The second thing we need to evaluate the magnitude of agglomeration benefits of transport investments is quantitative estimates of the returns to urban agglomeration. In other words, we require empirical verification of the existence and magnitude of the relationship between productivity and urban scale or density. Preferably, we want to examine this relationship separately for different sectors of the economy because we know that some sectors are more urbanised than others and thus likely to gain more from agglomeration.

To do so we estimate elasticities of productivity with respect to agglomeration, which we denote ρ , for defined sectors of the economy. In general, estimation requires two sources of data to proceed.

- i. **Representation of spatial variance in productivity** – this is usually achieved via spatially referenced firm level production data, on output and inputs, or spatially referenced wage data.
- ii. **Representation of spatial variance in agglomeration** – as mentioned previously, ED measures can be calculated at the zone level to represent agglomeration using appropriate measures of mass and impedance.

With these two sources of data we can proceed with econometric estimation of the agglomeration elasticity. Our objective is to estimate the effect of agglomeration on productivity.

With firm level data, productivity is represented in a model for Total Factor Productivity (TFP), which determines a statistical relationship between economic output (i.e. of firms, cities, regions etc.) and factor inputs (i.e. labour, capital, materials, etc.). This is done via a production function model which specified the nature of the relationship between the inputs that are used in production and the outputs that are produced. To evaluate the productivity effect of agglomeration, an ED measure is included in the production function as a shifter of productivity (ω) and an estimate the agglomeration elasticity:

$$\rho = \frac{\partial \log \omega}{\partial \log ED}.$$

With wage data, productivity is represented as labour productivity (LP). Models for LP typically make the assumption that factor inputs are paid the value of their marginal product, and they then take the wage rate (i.e. for workers, cities, regions etc) as an implicit measure of labour productivity. Estimating equations for wages can be derived from production functions by assuming optimising behaviour on the part of firms (i.e. profit maximisation or cost minimisation). Alternatively, at the level of individual workers, we can specify an equation where the wage (productivity) of some worker in a given location is explained by a set of worker-specific variables (education, age, skills etc.) and a set of ‘environment’ characteristics which include agglomeration economies. Again, the objective is to estimate the agglomeration elasticity, in this case given by

$$\rho = \frac{\partial \log w}{\partial \log ED},$$

where w is a measure of the wage used to represent productivity.

Whatever approach is adopted, there are three key issues shape the specific nature of the models to be estimated.

1. Sectoral breakdown – existing empirical evidence shows that the magnitude of agglomeration elasticities tends to vary across sectors of the economy. For this reason estimation of TFP or LP models should, as far as possible, be conducted separately for some reasonable disaggregation of economic sectors. This is important in evaluating the benefits of transport schemes as the local sectoral makeup of the economy can have an important bearing on how large any agglomeration benefits might be. The current elasticity estimates used in the UK provide a broad four sector disaggregation of the economy: manufacturing, construction, consumer services, and business services.

2. Data type – agglomeration elasticity values can be estimated using either cross-sectional or panel data. In a cross-section a sample of observational units are each observed at a single point in time. With panel data cross-sectional units are observed at different points in time. The vast majority of recently published empirical work in the field makes use of panel data, and the general rule of thumb in the literature is that panel data are superior for reliable estimation.

3. Data aggregation – as indicated above, the TFP and LP models could be estimated using data at various level of aggregation. Early studies of agglomeration tended to use aggregate production function or wage models for spatial zones, with agglomeration measured by city size. More recent studies tend to be based on microdata for firms and workers, typically in panel form. Micro panel data are generally regarded as superior for estimation of agglomeration elasticities. This is in part because they provide consistency with the underlying economic theory in the sense that they correspond with the micro-level at which behaviour is assumed (i.e. profit maximisation or cost minimisation), but also because they allow for application of more sophisticated econometric models.

Whether TFP or wage models are used, there are key methodological challenges that need to be addressed in estimating agglomeration elasticities to ensure that a *causal* estimate of transport improvements can be obtained. These are referred to in the literature as sources of endogeneity, which if not addressed adequately in the econometric models can lead to biased and inconsistent estimates of the agglomeration elasticity. Graham and Gibbons (2019) provide a comprehensive review of this literature. They show that substantial differences in the magnitude of estimates can arise due to difference in the methodological approaches used for estimation, and in particular, the manner in which studies have, or have not, attempted to correct for sources of 'endogeneity'. In general, failure to address endogeneity concerns adequately tends to cause an upward bias in elasticity estimates.

7.5.3. Elasticity values used to assess agglomeration benefits in the UK

The official agglomeration elasticity values used for appraisal of WEBs of transport schemes in the UK were estimated by Graham *et al.* (2009). Estimation made use of extensive firm level panel data for four broad sectors of the economy: manufacturing, construction, consumer services and business services. The analysis was based on a production function model with a control function approach to addresses potential sources of endogeneity and to allow for unobserved firm level heterogeneity.

Table 7.1. Summary of UK agglomeration elasticities

	<i>sic</i>	<i>agglomeration elasticity (ρ)</i>	<i>alpha</i>	<i>N</i>
Manufacturing	15-40	0.024 (0.002)	1.122 (0.127)	21,363
Construction	45	0.034 (0.003)	1.562 (0.159)	12,044
Consumer services	50-64	0.024 (0.003)	1.818 (0.190)	17,968
Business services	65-75	0.083 (0.007)	1.746 (0.144)	8,236
Economy (weight aver.)	15-75	0.044	1.659	

Notes: *sic* – standard industrial classification, *agglomeration elasticity* – elasticity of productivity w.r.t effective density (*ED*) (ρ as define above), *alpha* – the distance decay parameter of the *ED* metric, *N* – number of observations.

The key empirical results of their research are summarised in the table below. The *ED* measure used in this research is

$$ED_i = \sum_{j=1}^n \frac{E_j}{d_{ij}^{\alpha}},$$

where in addition to variables already defined, *E* is employment. The study estimates the elasticity of TFP with respect to *ED* (i.e. the agglomeration elasticity)

$$\rho = \frac{\partial \log \omega}{\partial \log ED}.$$

Note that a “distance decay” specification of effective density is used via the exponent α , which allows agglomeration externalities to diminish over distance. In addition to estimating the agglomeration elasticity, Graham *et al.* (2009) also estimate the value of $\hat{\alpha}$ using non-linear least squares. The motivation for identifying this parameter is that in assessing the agglomeration benefits of transport investments it is useful to understand the spatial scale over which these externalities are distributed.

The results from Graham *et al.* (2009) yield an overall agglomeration elasticity of 0.04 across all sectors of the economy. For manufacturing and consumer services the estimated elasticity is 0.02, for construction 0.03, and for business services 0.08. The distance decay parameter is found to be approximately 1.0 for manufacturing, but around 1.8 for consumer and business service sectors and 1.6 for construction. This implies that the effects of agglomeration diminish more rapidly with distance from source for service industries than for manufacturing. However, the relative impact of agglomeration on productivity is larger for services than it is for manufacturing.

A summary of agglomeration elasticity estimates from other recent studies of European countries is shown in **Table 7.2** below.

Table 7.2. Summary of recent European agglomeration elasticity estimates

Study	country	period	data	aggregation	Mean elas	Median elas
Ahlfeldt <i>et al.</i> (2015)	Germany	1936-2006	PD	regions	0.062	0.066
Brulhart & Mathys (2008)	Europe	1980-2003	PD	regions	0.080	0.055
Ciccone (2002)	Europe	1992	CS	regions	0.047	0.045
Combes <i>et al.</i> (2010)	France	1988	PD	worker	0.035	0.037
Combes <i>et al.</i> (2008)	France	1988	PD	zone	0.052	0.035
Combes <i>et al.</i> (2012)	France	1994-2002	PD	plant	0.090	0.070
Di Addario & Patacchini (2008)	Italy	1995-2002	PD	worker	0.010	0.010
Holl (2012)	Spain	1991-2005	PD	firm	0.089	0.047
Marrocu (2013)	Europe	1996-2007	CS	regions	0.036	0.041
Martin <i>et al.</i> (2011)	France	1996-2004	PD	plant	0.011	0.010
Mion & Naticchioni (2005)	Italy	1995	PD	worker	0.034	0.022

The empirical evidence from these studies supports the theory that agglomeration economies exist and that they induce higher productivity for firms and workers. However, there are differences in the magnitude of estimates reported. Mean values range from 0.01 to 0.09. The average of the means is 0.049, very close to the average UK value.

Melo *et al.* (2009) investigate why difference in estimates occur by conducting a meta-analysis. They find substantial differences in the magnitude of elasticity estimates across countries and industrial sectors, suggesting that responses to agglomeration are context and industry specific. In fact, elasticities estimates tend to be higher for business and producer services sectors; such as finance, insurance, real estate and professional services; than for manufacturing and construction. Melo *et al.* (2009) also find that substantial differences can arise due to the methodological approaches used for estimation, and this is evident both between and within studies. Of particular importance is the manner in which studies have, or have not, attempted to correct for

potential sources of 'endogeneity'. In general, failure to address endogeneity concerns adequately tends to cause an upward bias in elasticity estimates.

7.5.4. *Quantifying the agglomeration benefits arising from transport schemes*

Calculation of WEBs using agglomeration elasticities is relatively straightforward. To evaluate the agglomeration benefits of a transport scheme we first use a transport model (or transport and LUTI model) to calculate the level of agglomeration (static or dynamic) that will prevail after some transport intervention has been made. Note that in these calculations the change in agglomeration is measured using generalised cost based EDs , although the agglomeration elasticities are estimated using distance based EDs . This is the procedure recommended by the UK DfT, justified on the grounds that it avoids double counting the benefits of congestion relief. We denote EDs based on GC using ED^{GC} .

We then select an estimate of the elasticity of productivity with respect to agglomeration (ρ) (for each defined sector of the economy) and calculate the economic value of the agglomeration WEB impact of the transport scheme via

$$\Delta y(\Delta ED^{GC}) = y(ED_1^{GC}) - y(ED_0^{GC}) = \rho \left(\frac{ED_1^{GC} - ED_0^{GC}}{ED_0^{GC}} \right) y(ED_0^{GC}),$$

where $y(ED_0^{GC})$ is a measure of economic output (i.e. GDP) in the base year, ED_0^{GC} is effective density in the base year and ED_1^{GC} is the predicted value of effective after the transport scheme is in place.

Applied work on transport appraisal indicates that even relatively small impacts on TFP via agglomeration effects can give rise to large aggregate GDP benefits. Inclusion of agglomeration effects within CBA has often been found add between 10% and 30% to conventional user benefits. The table below shows provides examples of assessment for different types of schemes, showing the percentage of total benefits due to agglomeration.

Table 7.3. Assessment of agglomeration benefits for various schemes

<i>Mode</i>	<i>Scheme</i>	<i>% benefits agglom.</i>
Rail	CrossRail	19%
Rail	High Speed 2	18%
Rail	CrossRail 2	29%
Metro	Grand Paris Express	21%
Road	Multiple (UK RIS)	15%
Aviation	Heathrow 3rd Runway	15%
Multi	170 UK Schemes (Eddington, 2006)	≈20%

Notes: 1) % benefits agglomeration. is the share of total project benefits accounted for by WEBs of agglomeration, 2) the table is compiled from various schemes appraisals published in print or online, 3) Eddington (2006) refers to the Eddington Report which presents CBA data for 170 UK scheme across different modes.

It is worth noting from the table that agglomeration benefits are found across very different types of schemes. Agglomeration effects on productivity are induced not just by dense urban schemes such as CrossRail, CrossRail 2 and the Grand Paris Express; but also by the UK long distance high-speed rail (HS2) scheme. Appraisal work for the HS2 line in the UK did in fact find small agglomeration effects over long distances, largely because the limited capacity of the line simply did not induce sufficient change in *EDs* from long distance travel. Instead, the case for agglomeration benefits due to HS2 investment (18% in the table above) rested on local increases in *ED* due to the freeing up of local branch line capacity as a result of building the high-speed line.

The Eddington Report (Eddington, 2006) looked at 170 UK schemes of different types, across different modes, and of various sizes of investment. Eddington found that on average WEBs of agglomeration accounted for around 20% of total assessed benefits. It did not find strong evidence of mode specific differences in agglomeration effects, and also argued that the size of the investment was largely immaterial to the value of the resulting Benefit-Cost ratio: small and medium sized schemes can offer just as good value for money as large. Eddington also showed that inclusion of WEBs in CBA can help to define ‘strategic priorities’ for investment. In the case of the UK, these priorities were deemed to be urban networks, international gateways with surface access (ports and airports), and inter urban corridors.

Regarding this last priority, inter-urban corridors, there is a legitimate question over whether agglomeration effects can be transmitted over long distances in the same way they are over shorter distances. For instance, can an investment in HSR generate agglomeration effects across (rather than within) cities. As shown above, empirical work provides evidence consistent with the existence of agglomeration economies, but also generally indicates that the geographic scope of agglomeration externalities is relatively localized. However, there are in fact no obvious characteristics of the sources or mechanisms of agglomeration discussed in the literature that would limit their

generation over longer distances. Thus, most of the arguments linking transport to agglomeration could hold in long-distance case, in the sense that if spatial interactions between economic agents are made more efficient, we may expect increasing returns.

8. CASUAL INFERENCE FOR EX-POST EVALUATION OF TRANSPORT INTERVENTIONS

8.1. Introduction

Ex-post evaluation seeks to quantify the impacts that transport interventions have had on defined outcomes of interest. Such evaluations can be made in relation to almost any type of ‘intervention’ including physical projects such as the construction of a new link or scheme, or policy interventions such as the imposition of speed limits or changes in transport prices.

In contrast to ex-ante appraisal, ex-post evaluation is conducted after the fact, and it therefore relies more on actual than predicted changes, and on observed phenomena rather than theory, as the basis for calculations.

There are three main approaches commonly used for ex-post evaluation

1. **Ex-post CBA** - CBA calculations are made sometime after the project has been completed. Unlike ex-ante CBA, observed rather than predicted values may be available to quantify key changes induced by the intervention. Ex-post CBA can provide useful information both on the impacts of the project itself and on how well ex-ante CBA was able to predict the benefits and costs of the scheme.
2. **Case study / observation** - available data along with primary data collection, via surveys or interviews, can be used to build up a case study of impacts. This approach tends to be more qualitative in nature, typically seeking not to arrive at precise quantification of impacts, but instead provide some general indications of changes that have taken place following a transport intervention.
3. **Causal inference** - statistical models are applied to data observed before and after transport interventions have been made in an attempt to estimate impacts that were caused by the intervention. Identification of ‘causal’ rather than ‘associational’ effects is the over-riding objective of this approach. The statistical techniques used rely on empirical methods as well as economic theory and have assumptions and properties that must be met in order to generate valid causal inference on impacts.

In this Section we review quantitative approaches for ex post evaluation based on causal statistical modelling.

8.2. Causal inference and ex-post evaluation: terminology and key concepts

The emphasis on causality arises from the conviction that transport policy is fundamentally concerned with cause-effect relationships. Typically, the underlying aim of intervention is to attempt to shape future outcomes via public action. For decision makers the question of interest is: what impact, or outcome, will proposed interventions have.

Ex-post evaluation seeks to answer this question. By applying statistical models to historic data it attempts to capture key relationships of interest and evaluate the net effect of past interventions on defined outcomes. In so doing it is of first order importance that we obtain a causal interpretation from observed data. In other words, that we uncover the *actual* effect of the transport intervention, and not effects associated with other related or extraneous events.

In this section, we define some key concept and terminology that are central to the discussion of the causal ex- post evaluation approaches that follow.

8.2.1. Terminology and key concepts in causal inference

In ex-post evaluation we seek to quantify the effect of a transport *intervention* (or *treatment*) on *outcomes* of interest. This is done using data on a sample of observations where the *observations* or *units* could be transport links, geographic locations, spatial zones, households, people, firms or some other defined entities that we believe are affected by the intervention. In conducting our ex-post evaluation, we recognise that units of observation have *characteristics* that are relevant to the outcomes that we observe, and we measure these characteristics as *covariates*. Thus, in a typical ex-post set up we observe a triple of data describing outcomes, interventions and unit characteristics (or covariates). Below we provide some additional detail on these key concepts and define associated terminology.

First, is the concept of *intervention* and the related notion of ‘treatment’. A *transport intervention* refers to any planned changed made to the transport system. Interventions could involve physical projects such as the construction of a new link or scheme, extensions to existing networks, and capacity increases; or could involve non-physical policies such as price changes or regulations. In causal inference studies, interventions are often referred to as ‘treatments’. Treatment is a generic term for a random variable that provides some measure of an intervention. Treatment variables could be binary, multivalued or continuous. The table below gives examples of transport interventions classified as treatment variables.

Table 8.1. Transport interventions classified as treatment variables

Binary treatment	multi-valued treatment	Continuous treatment
tolled / untolled route	frequency of service	network capacity
presence of speed camera	speed limit	length of segregated route
20mph zone designation	cars per train	density of intersections
peak / off-peak	no. of O-D routes	accessibility
pedestrianized / unpedestrianized	no. of network nodes	tax / subsidy rates

A second key concept related to intervention is that of *treatment assignment* and the *assignment mechanism*. Assignment refers to the way in which a treatment is allocated. Across our sample of units, some will be affected by intervention and some will not. We refer to the former as treated units and to the latter as control units. If an assignment mechanism is random, then the probability of being assigned to the treatment is uniform across all units. That is, although we observe treated and control units in the data following assignment of the treatment, the probability of any unit being in either group prior to the treatment being assigned was uniform across the sample. On the other hand, if an assignment mechanism is non-random, then the probability of being assigned to the treatment varies across units in the sample. Transport interventions are almost always non-randomly assigned, being allocated to specific locations or links for pre-defined reasons. This simple fact is crucial in attempting to identify the impacts of interventions ex post.

The third key concept to clarify here is the *outcome(s)* of interest. In ex post evaluation outcomes refer to the dependent variables that are the object of ex-post evaluation. For instance, we might want to evaluate the effects of a transport intervention on the following outcomes: travel times, travel speeds, congestion, air quality or economic productivity. The objective of ex-post evaluation is to estimate the effects of treatments on outcomes.

The final concept to be defined here is that of *confounding*. As mentioned above, the observed data for ex post evaluation typically measures characteristics of units in the sample. These could include baseline (i.e. pre-intervention) traffic speeds, traffic flows, travel times, economic productivity, land use patterns and so on. Some of these characteristics may be confounders, and if so they have to be dealt with systematically in conducting ex post evaluation. A confounder is a random variable that has an influence on the outcome of interest, but that is also important in determining (potentially non-random) assignment to treatment. Confounders are important for ex-post evaluation because if they are not addressed adequately in the statistical model employed, then there can be serious difficulties in drawing valid causal inference about the impact of treatments.

8.3. The potential outcomes framework for causal ex-post evaluation

Using the concepts defined above we can now outline a general modelling framework for ex-post evaluation of transport interventions. The aim of this framework and associated methods is to quantify impacts that have occurred due to explicit intervention (or ‘treatment’), net of other potential ‘confounding’ influences. Transport interventions are characterized by non-random assignment, and we show here that causal methods based on potential outcomes are readily adaptable for ex-post evaluation.

The typical set up for ex post studies is one in which the data available for estimation are realisations of a random vector, $Z_i = (Y_i, D_i, X_i)$, where for our n units of observation, $i = (1, \dots, n)$, Y_i denotes a response (or outcome) of interest, D_i is the treatment (or exposure) received (i.e. some transport intervention), and X_i is a vector of unit characteristics that are potential confounders. Throughout this chapter we will discuss models and model specifications in relation to these three key elements of causal models: Y , D and X .

Treatment variables can be binary, multivalued, or continuous. As mentioned above, there are a variety of different ways in which transport interventions could be used to define a treatment variable. We could specify a binary treatment variable, with some units in the sample classed as affected by the treatment (i.e. “treated”) and others as unaffected (i.e. “control”). Or we could seek to measure the extent to which units in the sample have been affected by the treatment, and thus create a variable that distinguishes units by “dose” (i.e. by some measure of the magnitude of intervention).

Whatever treatment variable we construct, it should provide the basis upon which we can define a set of *potential outcomes* for each unit in the sample. Thus, for unit i and binary treatment $D \in \{0, 1\}$ we can define two potential outcomes defining treated and control status: $Y_i(0)$ if $D_i = 0$, and $Y_i(1)$ if $D_i = 1$. For multivalued or continuous treatment, we can define a potential outcome $Y_i(d)$ associated with each dose of treatment d , with $Y_i = \{Y_i(d): d \in D\}$ denoting the full set of potential outcomes.

Ultimately, the objective of our analysis is to estimate the effect of treatments on outcomes. A key problem for causal inference, however, is that the data available for estimation reveal only actual outcomes not the full set of potential outcomes (which never occur together). The key insight of the potential outcomes framework for causal inference is that is that we do not have to observe all potential outcomes to identify the causal effects of a treatment. Under certain conditions the actual outcomes are sufficient, even under non-random treatment assignments. We now show the logic underpinning identification of causal effects within this framework for inference.

Once we have a defined treatment variable, the key estimands of interest, or quantities that we want to evaluate ex post, are Average Potential Outcomes (APOs) and Average Treatment Effects (ATEs). An APO measures the average outcome that would occur if all units in the sample were assigned to some defined treatment status. The ATE is the difference between an APO for some particular treatment status of interest and the APO under some base scenario (often taken to be ‘untreated’ or control status). For binary treatments the relevant APOs are

$$\mu(1) = E[Y_i(1)] \text{ and } \mu(0) = E[Y_i(0)] ,$$

where $Y_i(1)$ is the outcome for unit i under treatment and $Y_i(0)$ is the outcome for unit i under control status. The binary ATE is defined as

$$\tau(1) = \mu(1) - \mu(0).$$

For continuous and multi-valued treatments the APO under some treatment level d is denoted

$$\mu(d) = E[Y_i(d)] ,$$

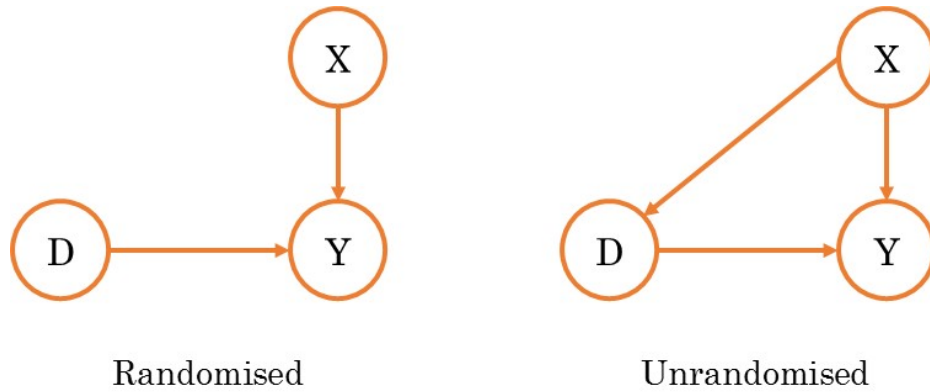
and the ATE takes the form

$$\tau(d) = E[Y_i(d)] - E[Y_i(0)],$$

for any dose of interest d .

Since we do not observe the full set of potential outcomes for units in the sample, one option to estimate APOs and ATEs would be to simply take mean outcomes for groups of units defined by treatment status and compare (i.e. in the binary case we could compare the mean outcome for treated units with the mean outcome for untreated units). However, in the absence of experimental conditions we cannot assume that the treatment is assigned randomly, and consequently, simple comparisons of mean responses across different treatment groups will not in general reveal a ‘causal’ effect due to potential for confounding. The implications of non-random assignment are illustrated in the figure below.

Figure 8.1: Directed Acyclic Graph of randomised and non-randomised treatment assignment



Note that under randomisation the characteristics of units in the sample, denoted X , have no influence on the treatment received (i.e. on D). Consequently, outcomes are unconditionally independent of the treatment assignment mechanism (i.e. all units in the sample have an equal probability of being assigned to treatment). Under randomisation a simple comparison of mean outcomes by treatment groups will yield valid inference about the ATE for the population. Under a non-randomised assignment, the allocation of the treatment depends on characteristics X , which are themselves important in determining outcome Y . This means that units differ systematically by treatment status in their base characteristics, and thus some part of the *association* between the treatment and the outcome could be attributed to X rather than D .

Transport interventions are typically assigned non-randomly. For example, they may be assigned to induce improvements in congested locations, in accessibility constrained locations, or in locations with poor economic performance. If so, the characteristics of the locations to which transport interventions have been assigned differ in some systematic ways from the general characteristics of locations that did not receive an intervention.

Under non-random assignment we refer to X as confounders and note that approximation of $\tau(1)$ or $\tau(d)$ using sample averages by treatment status may not reveal a ‘causal’ effect because the averages may differ regardless of treatment status due to the confounding effect of X on treatment assignment. The key issue for successful causal inference involves identifying and adjusting for confounding factors.

Broadly speaking there are two ways of doing this. First, through model-based adjustment for confounding, in which differences between units in characteristics X ,

are measured and included within a model to obtain marginal causal effects (i.e. ATE net of confounding influences from X). Second, by developing models which exploit sources of exogenous variance to obtain causal estimates without explicit representation of X in the model. In the following sections we describe three of the most commonly used statistical models for ex post evaluation.

8.4. Causal statistical models for ex-post evaluation

In the previous section we provided a brief introduction to the potential outcomes framework for causal inference. In this section we review three specific methods that can be used for ex post evaluation.

8.4.1. Outcome regression

Outcome regression (OR) models are perhaps the most commonly used approach to evaluate the impacts of transport interventions in the existing academic and consulting literature. The idea is to develop a regression model which relates the outcome of interest (i.e. some defined metric for ex post evaluation) to the treatment of interest, while adjusting for confounding via a set of measured covariates. For instance, in the context of linear regression an OR model could be specified as

$$y = T\tau + X\beta + \epsilon,$$

where y is the outcome of interest, T is treatment and X is a vector of covariates and ϵ is an error term with some assumed properties. The parameter τ is used to provide an estimate of the effect of the treatment on the outcome.

The OR approach can fail when the available data do not allow for confounding to be adequately addressed. In the recent empirical economics literature, a common solution applied in these situations is instrumental variables (IV) estimation. The IV estimator is well known and widely used in. The key principles of IV estimation are:

1. Find a set of instruments which are exogenous to the outcome but highly correlated with the treatment.
2. Use the instruments to enforce orthogonality between the error term and an instrument transformed design matrix.

8.4.2. Propensity score models

The next statistical approach to be considered is that of propensity score (PS) modelling. The key feature of PS based estimators is that they attempt to estimate the impacts of treatments on outcomes by explicitly modelling the treatment assignment mechanism itself. This is done in two steps.

1. **Estimation of propensities scores (PSs)** - PSs are calculated to capture relationships between treatment assignment and confounding effects. For some defined treatment; which can be binary, multi-valued or continuous; a PS model is specified which measures the probability of being assigned to some level of the treatment given background confounding characteristics. For unit i we denote the PS by

$$\pi_i = P(T|X_i, \alpha).$$

2. **Outcome model** - the estimated PSs are then used to adjust for confounding in a model that relates outcomes to treatments.

To provide an intuitive perspective on the principle underpinning PS models consider the following scenario in the binary treatment setting. Suppose a treatment has been randomly assigned such that all units that can potentially receive the treatment have an equal probability of doing so, i.e.

$$\pi_i = P(T|X_i) = P(T) . \quad \forall i$$

In that case, there is no confounding and therefore no systematic difference between the characteristics of units that are treated or untreated because X makes no difference to the probability of being assigned to treatment. Consequently, if we take the average outcome of interest for treated groups and subtract from it the average outcome for untreated units we will obtain an unbiased estimate of the treatment effect:

$$\tau = E[Y_i(1)] - E[Y_i(0)].$$

If the treatment is assigned non-randomly, and depends on unit characteristics X , then the probability of receiving treatment is non constant across the sample. Consequently, simple comparisons of averages by treatment status will not in general reveal a ‘causal’ effect because of confounding influences.

The PS attempts to quantify the conditional probability of being assigned to some level of the treatment given confounding characteristics. In other words, it attempts to quantify π_i for all units in the sample and these values will be non-uniform if confounding is present. PSs are calculated by estimating the relationship between T and X using a regression model

$$E[T_i|X_i] = \psi^1\{m(X_i, \alpha)\} ,$$

for known link function ψ , regression function $m()$, and unknown parameter vector α . The estimated parameters of this model are then used to construct propensity scores

$$\hat{\pi}_i = P(T|X_i; \hat{\alpha}).$$

Once we have the estimated PSs, we can then create a *pseudo-sample* that simulates random assignment by using the conditional probabilities to mimic the representation of units that would occur under randomisation. For instance, if we find that unit A has twice the probability of being treated as unit B due to its X characteristics, we can simply double the influence of B in the sample.

The PS is defined for binary, multi-valued or continuous treatments. Once the PS model has been estimated it can be used to form a number of different nonparametric and semiparametric estimators. A key advantage in using the PS is that it avoids the need to condition on a potentially high dimensional covariate vector (i.e. X) and it is this dimension reduction property that allows for effective implementation of flexible distribution free estimators.

Recent examples of the application of PS based models for evaluation of transport interventions are by Graham *et al.* (2014) and Li *et al.* (2013). Graham *et al.* (2014) develop a PS based methodology to evaluate the impacts of urban road capacity expansions (i.e. a continuous treatment) in US cities. They look at impacts on traffic volumes, traffic densities and congestions. They show that the PS method can be used to derive a range of different impact measures including flexible “dose-response” curves. Li *et al.* (2013) use propensity score matching to evaluate the effect of speed cameras (i.e. a binary treatment) on accidents.

8.4.3. Difference-in-differences (DID)

Differences-in-differences is a “before and after” treatment effect estimation approach that is applicable when the effect of treatment on units can be represented as a binary variable (i.e. “treated” or “untreated”). It can reveal impacts associated with exposure to an intervention relative to non-exposure (control), but it cannot tell us about impacts by scale or “dose” of intervention.

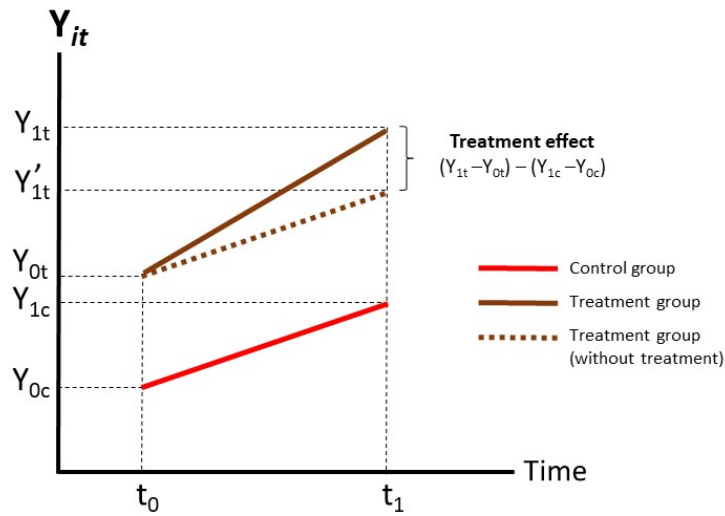
A problem in identifying treatment effects via OR models is that there may be unobserved differences between the treated and untreated units which affect outcomes and are also influential in treatment assignment (i.e. unobserved confounders). In addition, there may be temporal trends that affect the outcome variable due to events that are unrelated to the treatment. The DID approach provides an estimator for binary treatment effects which addresses such potential sources of bias by using information for both treated and control groups in both pre and post treatment periods.

The DID method works as follows. Observations are collected for two groups for two periods. One of the groups is the treatment group which is exposed to the treatment in one period. The other group is the control group which receives no exposure to treatment during both periods. The average change over time in the non-exposed (control) group is then subtracted from the average change over time in the exposed

(treatment) group to estimate the treatment effect. This double differencing, the so called “difference-in-difference” method, removes biases in the second period comparison between the treatment and control groups that could be the result of permanent differences between those groups, as well as biases from comparison over time in the treatment group that could be the result of time trends unrelated to the treatment.

The DID method is illustrated graphically below.

Figure 8.2: Graphic illustration of the difference-in-differences approach



In this diagram Y_{0c} refers to the average pre-treatment outcome for the control group, Y_{1c} refers to the average post-treatment outcome for the control group, Y_{0t} refers to the average pre-treatment outcome for the treatment group, and Y_{1t} refers to the average post-treatment outcome for the treatment group.

The DID approach estimates the treatment effect as

$$\tau = (Y_{1t} - Y_{0t}) - (Y_{1c} - Y_{0c}).$$

This can be achieved via linear regression. For instance, we can estimate the model

$$Y_{i,t} = \mu + X_i' \beta + \alpha T_{i,t} + \delta^* t + \tau T_{i,1} + \varepsilon_{i,t},$$

for units of observation i , $i = (1, \dots, N)$ in binary time periods $t \in \{0, 1\}$, with $t = 0$ representing the pre-treatment period and $t = 1$ the post-treatment period. In this model $T_{i,t}$ is the treatment indicator variable such that $T_{i,t} = 1$ if unit i has been exposed to the treatment prior to period t and $T_{i,t} = 0$ otherwise, δ is a time specific component, and $\varepsilon_{i,t}$ is a potentially autoregressive error with mean zero in each time period. The effect of the treatment is captured by the parameter τ .

In order to estimate the treatment effect, the DID method relies on the strong identifying assumption that both treated and control groups would have followed parallel paths over time. This is shown by the dotted line in **Figure 8.2** above. This line illustrates how the DID model predicts the average outcome for treated units that would have occurred if they had not received the treatment. It does so using information from the control group and by assuming that treated and controlled groups have a parallel trend. Adding covariates to the linear DID regression (i.e. X) can help in satisfying the parallel trend assumption because it is then assumed to hold conditional on those covariates, thus accommodating heterogeneity in outcome dynamics between the two groups.

The DID model provides a viable and popular approach for binary treatment effect estimation that has previously been used to evaluate ex-post impacts of transport interventions. While applicable only to binary treatments, it typically requires less data than the OR and PS models, although it is necessary that pre and post treatment data is available for both treated and controlled groups and it is also desirable that covariates are available to help strengthen the plausibility of the parallel trend assumption.

A recent example of the application of DID models for ex post evaluation is by Carbo *et al.* (2019), who study the impacts of the introduction of high-speed rail (HSR) between Madrid and Barcelona on a variety of economic outcomes. Using DID, they find that provinces with stops on the HSR line have experienced an increase of 2.4% in economic output, 3.3% in numbers of firms, and 1.1% in productivity. However, although they found evidence of growth in economic activity in provinces that received HSR stops, due to limitations of their data, they were not able to distinguish whether these gains represented net growth for the Spanish economy as a whole, or displacement from control provinces

9. GOVERNANCE AND INSTITUTIONAL DESIGN⁴⁴

9.1. The provision of transport infrastructures

The role of the public sector in the provision of transport infrastructures and the regulation of transport services is well-known. In the last decades, the liberalization of transport markets and the privatization of public companies have substantially increased the role of markets and private ownership in the functioning of the economic activity. Consequently, the public sector has diminished in terms of its size but not in importance.

Transport plays a crucial role in economic growth and sustainable development by allowing people access to health and education services, by increasing employment opportunities, by improving the exchange of information and, in general, by ensuring social cohesion. The transport sector also contributes to the reduction of costs of goods and services, and the increase of productivity through economies of scale and agglomeration. Transport infrastructures are characterized by high cost of investment, high proportion of sunk costs, demand uncertainty and irreversibility, but this does not justify by itself the provision of transport infrastructures by the public sector.

Other reasons, such as the different size and configuration of transport networks resulting from a public or private perspective, the design of contracts, the regulation of infrastructure monopolies resulting from the unbundling of railways or ports, the access to private operators to the basic monopolized infrastructure, justify the intervention of the public sector with the aim of aligning the interests of the different agents in the economy.

The provision of transport infrastructures consists of different stages such as planning, evaluation, decisions on how the private sector should participate in the construction, maintenance and/or operation of the transport services, and regulation in a broad sense (from price and quality regulation to contract enforcement).

The public sector must be in charge of the planning of the transport infrastructures for different reasons such as network design, to avoid duplication, or to ensure the construction of segments that are socially necessary but not profitable for private firms. However, the public sector does not have to be directly responsible for the construction, maintenance and/or operation of such transport infrastructures.

⁴⁴ This section draws on de Rus and Socorro (2010), Engel *et al.*, 2014, de Rus (2015), and de Rus and Socorro (2017).

There are three main approaches for the intervention of the public sector in the provision of transport infrastructures: public provision, public-private partnerships (PPPs), or privatization. Under public provision, a private firm builds the infrastructure in exchange of a negotiated payment and, once it is constructed and paid, it concludes its contractual agreement with the government. Then, the infrastructure is maintained and operated directly by the public sector. It is also possible that the maintenance and the operation of the infrastructure is performed by the private sector but through a different contract in which the private sector receives a payment for the tasks included in the contract without any risk or connection with the construction phase. On the contrary, through PPPs, the same private firm that constructs the transport infrastructure also operates and maintains such an infrastructure during a certain period. Therefore, the private firm contemplates the whole process with their trade-offs in terms of costs: a construction design can reduce maintenance costs, etc. Privatization differs from PPP in that the transport infrastructure is permanently transferred to the private firm.

In the provision of public infrastructure, the government must decide the type, the place and the timing of projects, and to design the contracts to be offered to the private sector for the construction of the transport infrastructure and the maintenance and operation (if applicable) during the life of the projects. There is an overwhelming body of evidence of government failure to deal with these problems (Engel *et al.*, 2014). The evidence shows that public intervention in the transport system is far from being optimal. Wasteful investment, inefficient pricing, poor regulation and a poorly designed private participation in the construction and operation phases of infrastructure provision have been common features in many countries all over the world. The common institutional design of the ministry of transport in many countries, with a separation of the modes of transport in different general directorates, contributes to the explanation of why transport networks are developed today in the way they do. With the present type of governance, it is perfectly possible to have a simultaneous and suboptimal expansion of the high-speed rail (HSR), roads and airport networks even in the case of mutually exclusive projects to address the same transport problem.

The evaluation of transport projects and the regulation of private participation and prices should be performed by independent agencies. The institutional infrastructure required by modern societies requires independent regulators. This is commonly accepted but it is also necessary an independent agency that guarantees the rigorous selection of investments. Then, private participation has to be based on auctions, and contracts designed in accordance with economic theory and best international practice. The objective is to maximize social welfare, by reducing the political interference in the technical phase of the project evaluation, and by guaranteeing the selection of good projects and its implementation by the most efficient firms.

The independent agencies must operate with total autonomy and independence with respect to firms and government, if we do not want them to become useless bureaucracies. Independent agencies are designed to be isolated from political pressures and are usually headed by a commission or a board of members chosen for their technical skills and reputation (and not for their political ideology). In this sense, the government should have limited power to choose or dismiss a member of an independent agency. Independent agencies should perform their tasks in an efficient, transparent, flexible and dynamic manner in order to maintain their social credibility. Since their origin is not democratic and their members do not depend on citizens' votes, their performance should be, in turn, assessed by independent evaluators in order to guarantee that they correctly fulfill their tasks and objectives.

9.2. The planning of transport infrastructures

The planning of transport infrastructures should be performed by the public sector in order to guarantee the mobility of all individuals across the country (even in those segments that are not profitable for private transport operators). In order to do so efficiently, the government should consider the overall transport network and all possible transport alternatives, avoiding the duplication of transport infrastructures to solve the same mobility problems or the investment in excess capacity. A clear example of duplication of transport infrastructures in Spain is the investment in HSR and airports. The investment in HSR has created the largest HSR network in Europe that in most cases overlaps the already existing network of airports to solve the same medium distance mobility problem. But, why governments may invest in transport infrastructures that are not optimal from the social point of view?

There are several reasons that explain why governments may deviate from a welfare-maximizing behaviour. The first is that the government tries to maximize the probability of re-election (Downs, 1957; Niskanen, 1971; Sobel, 1998; and Robinson and Torvik, 2005). In the case of transport investment, the typical project "...takes a little from a large group (e.g. all taxpayers) while hugely benefiting a small group (e.g. a specific subset of travellers). Since any single investment is politically rational in this way, over-investment is the expected outcome" (Mackie *et al.*, 2014). The second is the capture theory (Stigler, 1971) and/or the interest groups pressure to guide government policies to their advantage (Becker, 1983). In some cases, a third reason could be added, concerning the absence of economic principles in the public agencies of transport infrastructure. In many countries, the institutional design of the Ministry of Public Works has favoured the dominance of an engineering view in planning, disregarding the existing economic knowledge on economic evaluation and the problems of incentives for contract design and regulation in a context of asymmetric information, and different agents and objectives (see Engel *et al.*, 2014).

Efficient planning implies considering the overall transport network and all possible alternatives to solve a certain mobility problem. Therefore, transport infrastructure investment decisions cannot be taken within a Ministry of Public Works divided in different directorates by mode of transport that take investment decisions in isolation and without considering the important cross-effects between different modes. In contrast, a unique division should be in charge of the economic planning of the whole national transport network, avoiding duplication of transport infrastructures or, even worst, the configuration of a suboptimal transport network in the long-run. Indeed, the economic planning of transport infrastructure projects need to look carefully to the dynamic process associated with the initial decisions. The existence of multiple equilibria in the long-run to solve the same mobility problem and the possibility of ending up in a bad equilibrium when the evaluation concentrates on individual projects and loses the larger picture of the long-term intermodal effects needs to be highlighted. This is the problem of dealing with a project in isolation, ignoring relevant interactions with other markets and the dynamic process during the lifespan of the project.

The case of HSR versus air transport is illustrative in countries where both systems can be considered mutually exclusive. In countries with low population density, the usual base case is a network of airports with enough capacity to provide infrastructure for point-to-point medium distance trips. HSR infrastructure is a technology for high volume corridors, which is expensive and has a high proportion of sunk costs. The irreversibility of investment is one of its main characteristics. Compared to HSR, sunk costs and irreversibility are lower for the air transport infrastructure. The reason is twofold: On the one hand, the cost of building airports depends on the level of demand, since the higher the level of demand, the higher the size of the airport. On the contrary, the costliest part of the HSR infrastructure are the rail tracks and, thus, the cost of constructing the rail infrastructure varies little with the level of demand. On the other hand, once two regions have been connected with airports, only one more airport is needed to connect a third region (half of the previous investment). However, once two regions have been connected by rail, the cost of connecting the third region is almost the same (de Rus and Socorro, 2017).

Hence, unless the HSR project is carefully evaluated looking at all alternatives and the long-term consequences of the investment, it may well be that the initial decision of building HSR lines ends up with an undesirable equilibrium in which the wrong technology displaces a cheaper, more efficient, financially sustainable and reversible alternative. HSR is a technology to solve transport problems as well as air transport. Thus, technical neutrality has to be a key component in the planning process.

9.3. The evaluation of transport infrastructures

Transport infrastructure investment requires significant amounts of public funds, in a context of high proportion of sunk costs, irreversibility and demand uncertainty. In this context, a sound CBA is fundamental in order to guarantee that only projects that increase social welfare will be finally carried out. However, as already argued, there may be cases in which governments deviate from a welfare-maximizing behavior and invest in projects that are not optimal from the social point of view. If this is the case, and a CBA is required prior to the investment, the government has strong incentives to overestimate benefits and underestimate costs (Flyvbjerg, 2014), and the CBA becomes just a bureaucratic requirement.

A related and highly relevant subject concerning the economic evaluation and the institutional design appears when there are various levels of government and the government that evaluates the project is not the one that is financing the investment. In particular, suppose that transport projects are evaluated within a framework in which different levels of governments are implied (for example, national and regional governments, or supranational and national agencies), and where the objectives of the agents involved are not aligned. Assuming that a positive net present value is a requirement to get the project through, and the objective of a regional government (or a national government) is to get his project approved and totally financed by the national government (or partially funded by the supranational agency), the incentives to overestimate benefits and underestimate costs are obvious. This separation between who promotes and who pays the project also affects the decisions on infrastructure capacity and the kind of technology to be chosen (de Rus and Socorro, 2010).

To avoid all these inconveniences, an independent agency, with sufficient means and unquestionable technical reputation, should evaluate and prioritize all infrastructure investments. In the economic evaluation of transport infrastructures, all alternatives should be included (which implies including not only all possible transport modes to solve the mobility problem but also the possibility of postponing the investment), and all costs should be considered (independently of who is paying and the different levels of governments involved). Ideally, this evaluation agency should evaluate all public projects and not only transport projects since, given governments' budget constraint, all public projects compete with each other. Independent evaluation and formal approval of infrastructure projects is needed in order to achieve efficient investment decisions.

For sufficiently large projects, additional improvements are possible if, after the social evaluation stage, the project is reviewed by a high-level board with the authority to approve it, to ask for modifications, to postpone or to reject the project. This board should go beyond the CBA and assess the overall feasibility of the project, its

interactions with other projects, the sources of financing, etc. To perform its role, the board should be politically and financially independent and its members should be chosen for their professional skills and reputation and should bear responsibility for their decisions (Engel *et al.*, 2014).

9.4. The regulation of private participation in the provision of transport infrastructures

It is the government's responsibility to plan the infrastructure network, but it does not make much sense to build, or exploit, them directly as monopolists in the case of services that can be offered by private firms competing in the market. If competition in the market is not possible, competition for the market should be introduced, by awarding a temporary monopoly to the best bidder chosen through a well-designed bidding mechanism and contract design. Once the contract has been designed, it should be monitored in order to guarantee that all terms of the contract are fulfilled and avoid renegotiation.

Contract design should be analyzed in a context of asymmetric information, considering moral hazard and/or adverse selection problems. In contract theory, the difference between moral hazard and adverse selection is related to the moment in which the asymmetric information appears, either before signing the contract (adverse selection problem) or after the contract is signed (moral hazard problem). Contract design should be performed in order to provide the right incentives to private firms and look for the optimal allocation of risk. Quality should be also controlled in order to guarantee that private firms do not decide to cut costs by reducing quality. If the quality is verifiable, it can be included in the contract. If the quality is not verifiable and, thus, it cannot be contractible, the regulator should give private firms the right incentives to provide the socially desired level of quality.

The economic theory of procurement uses mechanism design to analyse the problem of asymmetric information and incentives in principal-agent models (see Laffont and Tirole, 1993, for a review of this literature). Although most theoretical papers in the literature suggest a menu of contracts to deal with asymmetric information problems, this menu of contracts is rarely used in practice. Instead, the vast majority of contracts are variants of simple fixed-price and cost-plus contracts. In fixed-price contracts, the contractor is offered a fixed quantity for completing the project. A cost-plus contract does not specify a fixed price, but rather reimburses the contractor for the costs of the project. Fixed-price contracts have some advantages and disadvantages over cost-plus contracts. On the one hand, fixed-price contracts provide the strongest incentives for cost reduction. On the other hand, fixed-price contracts leave all risk with the contractor, increasing the probability of renegotiation. On the contrary, cost-plus contracts suppose no risk for the contractor but no incentives to reduce costs.

When there are different levels of governments and supranational (national) governments finance national (regional) governments, national (regional) governments may lose their incentives to apply optimal pricing and offer optimal contracts to the private sector. In this case, we have to distinguish two different levels regarding national infrastructure investment. The first level is related to the institutional design in which supranational (national) funds are obtained, that is, it focuses on the supranational (national) planner and national (regional) government relationship.

Once supranational (national) funds are obtained, the second level is related to the selection of the contract to be used for the construction, maintenance and operation of the infrastructure. Thus, the second level analyzes the national (regional) government and the contractor relationship. In an asymmetric information framework, the characteristics of the financing mechanism may be crucial in terms of incentives to select an efficient contract for the construction, maintenance and operation of the infrastructure in the second level, that is, a contract looking for the minimum costs and the maximum revenues from the social point of view.

In order to illustrate this, consider three kind of supranational funding mechanisms used in reality: total cost-plus, sunk cost-plus, and fixed-price financing mechanisms. With a total cost-plus financing mechanism, a percentage of the difference between investments costs and the discounted profits of the project are covered by the supranational organization. Thus, both investment costs and maintenance and operating costs may be financed. On the contrary, with a sunk cost-plus financing mechanism, the supranational organization covers just a percentage of the difference between investments costs and the discounted revenues net of maintenance and operating costs. Thus, in this case only sunk costs may be financed. The higher the total amount to be co-financed the higher are the total investment costs and the lower are the revenues, so it is a kind of cost-plus financing mechanism which penalizes the revenue generating projects. Finally, with a fixed-price financing mechanism, countries are granted with a fixed quantity to support the costs of the project chosen by the country.

With a total cost-plus financing mechanism, and assuming there is no congestion, the incentive is to give free access to the new infrastructure, the market quantity is excessive, and the level of supranational financing disproportionate. In contrast, with a sunk cost-plus financing mechanism the socially optimal price is always implemented, though there is no incentive in being efficient. Finally, with a fixed-price financing mechanism the maximal efficiency may be achieved, and the socially optimal price is always implemented.

For the regulation of private participation, an independent agency, different from the one in charge of the economic evaluation of projects, should be created. This

independent agency would be responsible for designing, awarding and managing concession contracts for private participation in all modes of transport, providing the appropriate incentives. This agency, in turn, will be composed of three independent units, the first one devoted to the contract design and awarding process, the second one devoted to the compliance of the contract, and the third one devoted to conflict resolution and evaluate possible renegotiations.

Tenders and contracts should be designed by a specialized and independent unit within the regulation agency. After a private firm or concessionaire wins the contract, compliance is important and independent supervision is required. Independency is essential in order to break the conflict between promotion of new investment and strict supervision of contracts. Finally, there should be a formal mechanism for conflict resolution. Conflicts between the concessionaire and the supervision unit should be arbitrated by an independent panel of experts. The panel should also review contract renegotiations, basing its decisions on technical, judicial and economic considerations. Cost risks should be borne by the concessionaire, as long as the contractual service standards remain unchanged. However, if additional investments or improved service standards are needed, the concessionaire has to be compensated. This requires that any renegotiation should be subject to independent review.

9.5. Price regulation

In **Section 4**, we discuss the importance of the pricing scheme in the investment decision. Moreover, we argue that pricing decisions affect differently the social welfare of alternative transport modes. Thus, when comparing different transport alternatives, a particular charging scheme may favor the creation of a particular transport infrastructure network, leading to long-term equilibria that would not be optimal under other charging schemes (de Rus and Socorro, 2019).

Under public provision, user fees are usually set too low for political reasons, either because politicians are more concerned in maximizing the probability of re-election rather than maximizing social welfare or due to the pressure of certain lobbies or interest groups. Pricing below the socially optimal price is inefficient, leading to overuse of the transport infrastructure, problems of congestion in case of scarce capacity, or lack of maintenance since user fees may not even cover maintenance and operating costs. On the contrary, under PPPs or privatization, prices might be too high and, thus, they should be regulated in order to guarantee that they maximize social welfare and not only private profits. The increase of private participation must be accompanied by an improvement in price regulation, a task that is the responsibility of the regulators in advanced countries. This is one of the challenges that the government has with respect to private participation in industries with limited competition: to obtain the best of the private participation in the benefit of the whole society.

The public agency responsible of pricing has to consider the overall transport network and the cross-effects between transport modes. The pricing scheme should be the one that maximizes social welfare, and not the one that maximizes the probability of re-election, or the welfare of certain lobbies or interest groups. Since the objective function of the government that promotes the project may be different from social welfare, this regulation agency must operate with independence from firms and the government.

In general, the economic evaluation of a transport infrastructure project requires optimal access pricing. If the base case implies suboptimal socially access pricing, the positive net present value of the investment is not a sufficient condition for implementing the project. In a first best world, the necessary and sufficient condition implies a positive difference in social welfare between the cases in which the new infrastructure is and is not constructed, with socially optimal access pricing being applied in both cases. This is not a result derived from the presence of uncertainty and irreversibility but from the interaction of access pricing and investment decisions and the need to consider as a benchmark the case in which social welfare is maximized, that is, the case in which the infrastructure is not constructed and socially optimal access pricing is considered.

The decision on the socially optimal access price to be charged for the use of a particular infrastructure must be taken by considering the existence of substitutions or complementarities between facilities (de Rus and Socorro, 2014). This result has important implication in terms of the institutional design of public agencies such as the Ministry of Public Works in many countries, where the division of management units is usually based on technological characteristics (road, air or rail) with charging decisions taken in isolation, based on the specific characteristics of the transport infrastructure and without considering the overall picture and the important cross-effects between different modes of transport.

Moreover, in the case of railways with vertical separation, the rail track administrator has to charge for the use of the infrastructure to the rail operator(s). The common objective is to achieve allocative efficiency with the double objective of optimally allocating the limited capacity of train lines between different services and operators; and deciding the optimal level of investment. Laffont *et al.* (1998) suggest different ways to fix access charges with the aim of maintaining the infrastructure and keeping the incentive to invest.

In the case of a monopolist service operator, the access charge should follow the traditional two-part tariff, with a fixed charge (e.g., annually) aimed to the coverage (partially or totally) of the fixed costs of the infrastructure and a variable charge per use (train or service) to cover the marginal cost.

In the case of several operators competing in the provision of rail services (the situation with HSR services in Spain in 2020), the access charge should be proportional to the use. To promote competition and facilitate entrants to operate in a market dominated by the incumbent (the traditional public monopoly) a fixed charge will discourage smaller entrants to compete, at least until they reach a size compatible with the payment of the fixed part of the tariff. Hence, the idea is to use a linear access charge, a kind of average pricing, higher than the marginal costs of the services, to contribute to the coverage of the infrastructure fixed costs.

ANNEX A. REVIEW OF SOME OFFICIAL GUIDELINES FOR COST-BENEFIT ANALYSIS IN TRANSPORT

There are different procedures for carrying out the economic evaluation of transport projects. Although most of them share the same principles, they are not always applied homogeneously. This makes it necessary for government decision-makers and CBA practitioners to resort to clear references and explicit rules, not only consistent with conceptual advances in economic theory, but also with practical lessons learned from past experiences, without ignoring in any case that evaluation processes in the real world are generally performed within limited information contexts.

Thus, both as motivation and overall reference for the rest of this document, this annex reviews some of the most recent official guidelines related with the methodology for the economic evaluation of transport projects. To this end, a selective search has been conducted for manuals, handbooks and other similar guides on this subject published in the last decade either by international agencies or national governments. We have finally selected the five documents summarized in **Tables A.1 to A.5**, all providing instructions for the general assessment of transport projects, with or without additional mode-specific recommendations. For each of these documents we have analysed their structures and compared their contents. Most of them are extremely practical, although they do not always renounce to discuss theoretical aspects of cost-benefit analysis. In some of these manuals the main text is also completed with annexes that develop technical details of CBA, or even provide examples or model worksheets. However, the issues that have been particularly covered in this review refer to the main aspects related to project definition, objectives and comparison of alternatives, the identification of benefits and costs, their measurement and the decision criteria regarding transport projects.⁴⁵

⁴⁵ To simplify internal references within this Annex, from now on, each of the reviewed manuals have been given a short name, as described in the first row of each of the records in the following tables.

TABLE A.1. Recent official guidelines for the assessment of transport projects (I)	
Short name	EU2015
Full name	<i>GUIDE TO COST-BENEFIT ANALYSIS OF INVESTMENT PROJECTS</i> ⁴⁶ <i>Economic Appraisal tool for Cohesion Policy 2014-2020</i>
Organization	European Commission Directorate-General for Regional and Urban Policy
Publishing date	June 2015 (previous version, 2008)
Language	English
Brief description	<p>An example of a standard set of guidelines for cost-benefit analysis in general, and for transport (and other sectors) in particular provided by an international institution.</p> <p>It updates and expands the previous edition of 2008 by including several considerations about the recent developments in EU policies and other international best practices. The document (364 pages) offers practical guidance on major project appraisals, as embodied in the European cohesion policy legislation for 2014-2020, and should be seen “as a contribution to a shared European-wide evaluation culture in the field of project appraisal”.</p> <p>Its main objective is to illustrate common principles and rules for application of the CBA approach into the practice of different sectors. The guide targets a wide range of users, including desk officers in the European Commission, civil servants in the Member States and in candidate countries, staff of financial institutions and consultants involved in the preparation or evaluation of investment projects. The text is relatively self-contained and does not require a specific background in project assessment. Its main part is structured into five main topics:</p> <ol style="list-style-type: none"> 1. <i>Project definition</i> 2. <i>Technical feasibility and environmental sustainability</i> 3. <i>Financial analysis</i> 4. <i>Economic analysis</i> 5. <i>Risk assessment</i> <p>The part devoted to transport projects (pages 77-144) also covers these topics, additionally addressing detailed issues on <i>traffic forecasting</i>, <i>travel time savings</i>, <i>accidents and other externalities</i>. This part also includes three case studies on road, rail and urban transport. Other concepts (on <i>social discount rates</i>, <i>shadow wages</i>, <i>willingness to pay</i> and <i>probabilistic analysis</i>) are discussed in the annexes.</p>
Source	https://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/cba_guide.pdf

⁴⁶ An academic paper that reviews in detail this document and provides further references to EU investment policies and regulations was published by Florio *et al.* (2018).

TABLE A.2. Recent official guidelines for the assessment of transport projects (II)	
Short name	EIB2013
Full name	<i>THE ECONOMIC APPRAISAL OF INVESTMENT PROJECTS AT THE EIB</i>
Organization	European Investment Bank Projects Directorate
Publishing date	2013
Language	English
Brief description	<p>Another example of general and sector-specific CBA guidelines provided by an international institution which refers itself as “the lending arm of the European Union”, although it is also active in non-Member countries.</p> <p>This guide (224 pages) was prepared by staff members involved in project appraisal and economic analysis, a procedure that plays a central role in EIB operations because it allows to judge whether an investment project will contribute to the economic growth and cohesion of the EU and the economic progress of its partners. The document combines standard economic appraisal techniques, including cost-benefit analysis, cost-effectiveness analysis and, more recently, multi-criteria analysis, taking into account the specific characteristics of each sector. It is not intended as a manual, nor as a textbook; it rather describes “how the EIB does it,” giving the general reader an overview of the methods used, and the specialist a guide to the application of different analytical tools across sectors.</p> <p>In the cross-sector methodological part, the document covers as main topics:</p> <ol style="list-style-type: none"> <i>1. The need for financial and economic appraisal.</i> <i>2. Definition of the counterfactual scenario.</i> <i>3. Incorporating environmental externalities.</i> <i>4. Wider economic impacts.</i> <i>5. Economic life and residual value.</i> <i>6. Social discount rate.</i> <i>7. Risk analysis and uncertainty.</i> <p>With regard to transport-specific issues, this guide additionally discusses, among others, topics on the <i>value of time, transport safety, vehicle operating costs and traffic categories</i>, with case studies on interurban railways, roads, urban public transport, airports and seaports.</p>
Source	https://www.eib.org/en/publications/economic-appraisal-of-investment-projects

TABLE A.3. Recent official guidelines for the assessment of transport projects (III)	
Short name	TAG2013
Full name	TRANSPORT ANALYSIS GUIDANCE
Organization	Department of Transport (United Kingdom)
Publishing date	2013 (last updated, may 2019)
Language	English
Brief description	<p>An example of detailed administrative procedures for transport CBA at the national level.⁴⁷ TAG is a set of public (web-based) instructions to conduct transport studies. It provides links to advice on how to set objectives and identify problems, develop potential solutions, create an appraisal model and how to conduct an assessment which meets the department's requirements to make a decision in a later stage. The web-TAG is divided into three levels according to the administrative decision process: for the senior responsible officer, for the technical project manager and for the appraisal practitioner (including modelling principles, reference data and workbook examples). Its basic structure can be summarized into the main following sub-documents:</p> <ol style="list-style-type: none"> 1. <i>The transport appraisal process</i> [General overview] 2. Unit A1-1. <i>Cost-benefit analysis</i> [Project definition, measurement] 3. Unit A1-2. <i>Scheme costs</i> [Cost measurement, uncertainty] 4. Unit A1-3. <i>User and provider impacts</i> [Consumer surplus, time savings, operating costs, reliability] 5. Unit A2-1. <i>Wider economic impacts appraisal</i> [Definition and measurement] 6. Unit A2-2. <i>Induced investments</i> [Indirect effects in other markets] 7. Unit A2-3. <i>Employment effects</i> [Indirect effects in labour markets] 8. Unit A2-4. <i>Productivity impacts</i> [Agglomeration and similar effects] 9. Unit A3. <i>Environmental impact appraisal</i> [Includes noise, air quality, greenhouse gases, landscape and water impacts] 10. Unit A4-1. <i>Social impact appraisal</i> [Accidents and security, accessibility and affordability issues] 11. Unit A4-2. <i>Distributional impact appraisal</i> [Equity issues] <p>There are additional, more specific sub-documents on modal CBA (for walking and cycling projects, aviation, rail and highway interventions and on the use of marginal external congestion costs to estimate decongestion benefits resulting from mode switch away from car use. Detailed <i>modelling principles</i> and <i>forecasting techniques</i> as well as <i>data sources</i> and references are also given.</p>
Source	https://www.gov.uk/guidance/transport-analysis-guidance-webtag

⁴⁷ The TAG is the practical application of the underlying regulation defined at the corresponding *Green Books* on project appraisal and evaluation. Although some of the references are more recent (HM Treasury, 2018), we will maintain the TAG2013 reference for consistency purposes.

TABLE A.4. Recent official guidelines for the assessment of transport projects (IV)	
Short name	ADB2013
Full name	<i>COST-BENEFIT ANALYSIS FOR DEVELOPMENT: A PRACTICAL GUIDE</i>
Organization	Asian Development Bank
Publishing date	2013 (previous version, 1997)
Language	English
Brief description	<p>A third example of general and transport-specific CBA guidelines from a non-European perspective, with a special focus on developing countries.</p> <p>This guide (395 pages) was written in response to the findings of an internal review of ADB procedures for project assessment. The heterogeneity of previous results across different sectors suggested the need for a technical harmonization accompanied by clear and practical examples. For this reason, the document includes not only a general methodological part, but also extensive and detailed case studies focusing on infrastructure (including transport).</p> <p>Thus, the first part focuses on the following issues:</p> <ol style="list-style-type: none"> 1. <i>Income distribution and poverty.</i> 2. <i>Risk and uncertainty.</i> 3. <i>Valuation issues.</i> 4. <i>Environmental sustainability.</i> 5. <i>Choice of social discount rate.</i> <p>On the other hand, the transport-related chapter (pages 225-276) includes a case study (road improvement) and covers the following topics:</p> <ol style="list-style-type: none"> 1. <i>Project definition.</i> 2. <i>Demand forecasting.</i> 3. <i>Project benefits.</i> 4. <i>Shadow pricing.</i> 5. <i>Developmental benefits.</i>
Source	https://www.adb.org/documents/cost-benefit-analysis-development-practical-guide

TABLE A.5. Recent official guidelines for the assessment of transport projects (V)	
Short name	CEDEX2010
Full name	<i>ECONOMIC EVALUATION OF TRANSPORT PROJECTS. GUIDELINES</i>
Organization	Ministerio de Fomento ⁴⁸ (Spain) Centro de Estudios y Experimentación de Obras Públicas (CEDEX)
Publishing date	2010
Language	Spanish, English
Brief description	<p>A final example of a transport-specific CBA manual, written from an academic perspective for its use as a general reference in its assessment processes both by the Spanish Ministry of Fomento (www.fomento.gob.es/) and its technical unit, CEDEX (www.cedex.es).⁴⁹</p> <p>These guidelines (142 pages) were intended for economists, engineers and public servants interested in transport projects, assuming that any potential user has a basic knowledge of economics. Its main objective is to provide clear rules and economic principles that help to identify and quantify the precise contribution to social welfare of any transport project, avoiding double counting and relevant omissions. The guide departs from the idea that public investment in transport should be assessed using cost-benefit analysis, and that the existence of opportunity costs for social resources implies that society as a whole should always consider whether what it gains from the project exceeds what it might have obtained allocating the same resources to alternative uses.</p> <p>The document is structured in five main sections:</p> <ol style="list-style-type: none"> 1. <i>Project definition.</i> 2. <i>Decision criteria.</i> 3. <i>Calculation of changes in social welfare.</i> 4. <i>Identification of social costs and benefits.</i> 5. <i>Measurement of social costs and benefits.</i> <p>The annexes discuss <i>demand forecasting issues, valuation of external effects, equity and territorial impacts, and institutional design</i>. A set of supplementary working papers also provide further discussions on <i>indirect effects</i> and <i>cost estimations</i>, with case studies and reference values.</p>
Source	www.evaluaciondeproyectos.es

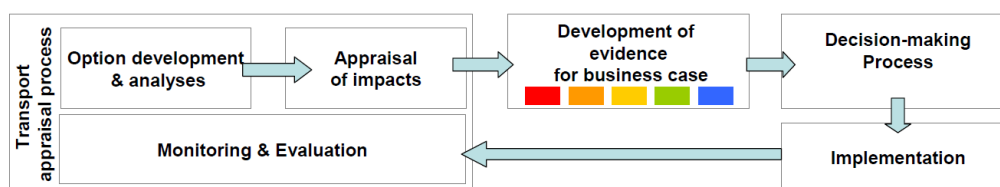
⁴⁸ Recently, in January 2020, the Ministerio de Fomento (Spanish Ministry of Development) changed its name to Ministerio de Transportes, Movilidad y Agenda Urbana. However, we will use the original name throughout the document, since that has been its official name during most of the period discussed in this report.

⁴⁹ The manual, and its accompanying documents, was one of the main results of a competitive research project, PT-2007-001-02IAPP, funded by CEDEX.

A.1. Project definition

All the five manuals summarised in **Tables A.1 to A.5** stress the importance of a good project definition to clarify the objectives of its assessment and its relevance within the wider and more general goals and policies of the relevant decision-maker. In **EU2015**, for example, CBA is explicitly required, among other elements, as a basis for decision on the co-financing of *major* EU projects (above 50 million euros). In all cases, both at national and supranational levels, the projects must be in accordance with sectoral and general regulations and reflect the strategic priorities of the corresponding project planning institutions. In fact, the decision process always involves at least one technical stage and a political stage, where the final approval is given (or not), regardless the assessment results.

Figure A.1. Project appraisal and decision process in the UK



Source: TAG2013

For example, transport project appraisal and the decision-making process in the United Kingdom, as described in **TAG2013**, are connected by a circular algorithm illustrated in **Figure A.1**, where *ex ante* (before the decision) and *ex post* (after the implementation) monitoring and evaluation procedures are explicitly taken into account. In the European Union and the other international institutions, the process is more complex due to specific institutional and political arrangements, but basically covers similar stages.

In general, the technical phase (the most relevant one for this document) of project definition always includes a presentation of the socio-economic context, as well as the definition of its objectives in terms of both the need of assessment and the project relevance. A clear definition of the project objectives is always necessary to identify the effects to be further evaluated in the CBA. The identification of effects should be linked to the project's objectives in order to measure the impact on welfare. "The clearer the definition of the objectives, the easier the identification of the project and its effects", as explicitly stated in **EU2015**.

In **CEDEX2010**, the definition of any transport project, considered as any external intervention in a transport market, includes in addition to this initial diagnosis, the enumeration of the alternatives to address the problem, and the choice of the baseline case against which the comparison is made.

In **EU2015** and **EIB2013** a great emphasis is given to technical and environmental sustainability, which are among the main elements of information to be provided in the funding request for major projects. In particular, a consistent demand analysis, options analysis, environment and climate change considerations and the technical design, cost estimates and implementation schedule are compulsory. For **ADB2013**, as expected, social development issues and impact on poverty are also relevant.

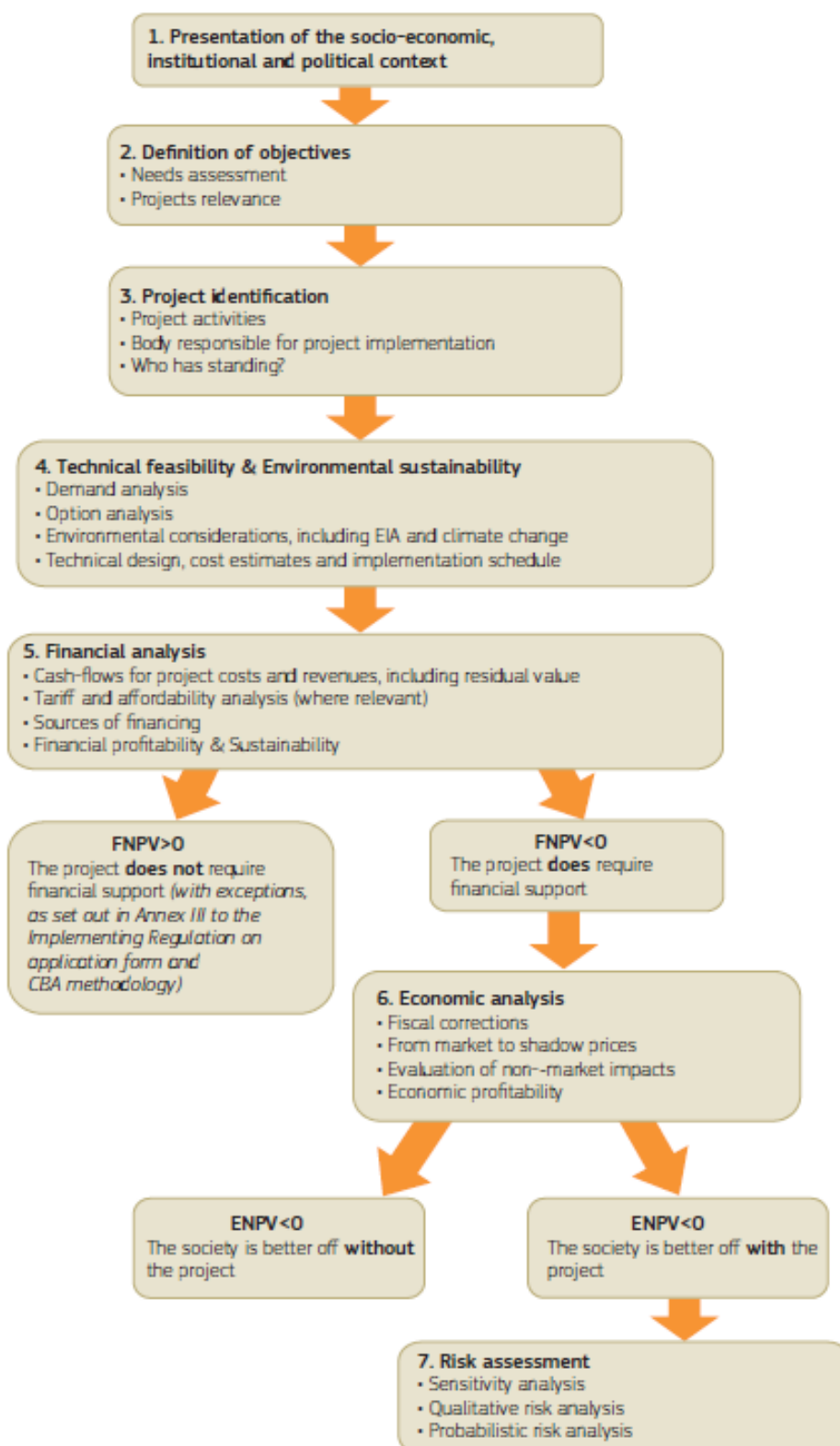
In all cases, CBA is performed using an incremental approach, by comparing a scenario ‘with-the-project’ with a counterfactual baseline scenario ‘without-the-project’. **EU2013** clearly states that when “...a project consists of a completely new asset, the without-the-project scenario is one with no operations. In cases of investments aimed at improving an existing facility, it should include the costs and the revenues/benefits to operate and maintain the service at a level that it is still operable (‘business as usual’) or even small adaptation investments that were programmed to take place anyway”.

Similarly, and depending on each case, in **EIB2013**, three basic types of counterfactual scenarios are identified against which to compare the project, ‘do-nothing’, ‘do-minimum’ (used by default) and ‘do-something-else’ (which would consist of an alternative approach to meet the objectives pursued by the project, and becomes an appropriate counterfactual for analysing project options, timing or phasing, once it has been recognised that “something” must be done). In all manuals, as explicitly stated in **CEDEX2010**, the base case is always a reference point used to compare what would have happened without the project, so it cannot have a static character but must incorporate what the evolution of the markets affected by the project would have been if the project had not been carried out.

A.2. Methods to identify benefits and costs

In addition to the necessary identification of the effects in connection with the project objectives, each of the reviewed manuals propose different approaches to how benefits and costs should be incorporated into the assessment. For example, in **EU2015**, the standard CBA is structured in seven steps, distinguishing between the financial and economic analysis.

Figure A.2. Project appraisal and decision process in the EU



Source: EU2015.

Thus, feasibility and sustainability are among the elements of information to be provided in the funding request for major EU projects. Although both analyses are not formally part of the CBA, their results must be concisely reported and include: a demand analysis (which identifies the need for an investment by assessing current and future demand), an analysis of the options (establishing and comparing a list of alternatives), considerations about environment and climate change,⁵⁰ and relevant issues regarding technical design, cost estimates and implementation schedules.

EIB2013 also has an explicit section on incorporating environmental externalities, stating that “external costs need to be added alongside operating and maintenance costs over the economic lifetime of the asset. This requires an estimate of the volume of externality and an appropriate unit price, or marginal external cost estimate (...)”. These references are provided in the manual, together with the main methodologies to estimate environmental impacts, although without further disaggregation by transport modes. Similar detailed concerns about environmental effects are found in **TAG2013** and **ADB2013**.

CEDEX2010 discusses different methodologies for identifying the effects of a transport project on social welfare. Two equivalent approaches, the change in agents’ surpluses and the change in the willingness to pay (WTP) and use of resources, are compared. Either one or the other approach must be used in each case, since their combination is not theoretically correct and a major source of practical mistakes.

EU2015 favours the WTP approach although suggests the use of a stakeholders’ matrix to avoid double-counting and/or missing relevant agents. On the other hand, both **ADB2013** and **TAG2013** seem to prefer addressing the calculation of users’ benefits using the conventional producer surplus and consumer surplus theory “where consumer surplus is defined as the benefit which a consumer enjoys, in excess of the costs which he or she perceives”, incorporating changes in travel time, changes in user charges (including fares, tariffs and tolls), and changes in vehicle operating costs met by the user (i.e. for private transport). **ADB2013** also distinguishes across agents, separately considering the surpluses of workers, investors and government. Following its poverty impact concern, this guide particularly favours the use of relative weights in adding surpluses to obtain the social welfare function. These guidelines suggest the importance of presenting how a project impacts on different groups (e.g. car users, public transport users, taxpayers), rather than hiding distributional impacts in the aggregation of resource costs and benefits. In fact, in some of the transport case studies presented in **EIB2013** (on roads), this is the suggested approach.

With regard to indirect effects and wider economic impacts, **EU2015** considers that, when direct effects are properly measured through shadow prices and all externalities have been fully monetised “indirect effects occurring in secondary markets (e.g.

⁵⁰ Although it is not a CBA manual, the European *Handbook on the External Costs of Transport*, published in 2019 and widely discussed in **Section 6**, also provides valuable information and reference values for the economic assessment of transport projects from the point of view of environmental externalities.

impacts on the tourism industry) should not be included in the evaluation of the project's costs and benefits. The main reason for not including indirect effects is not because they are more difficult to identify and quantify than direct effects, but because - if the secondary markets are efficient - they are irrelevant in a general equilibrium setting, as they are already captured by the shadow prices. Adding these effects to the costs and benefits already measured in primary markets usually results in double-counting”.

EIB2013 shares an equally cautious view, suggesting that when the benefits of a project, measured using standard appraisal techniques, fail to outweigh the costs, it may be tempting for promoters to search for wider economic impacts. In line with standard CBA practice, they state that the central focus of the economic appraisal is to capture accurately the flows on relevant primary markets, and there is an overall presumption against including impacts on secondary markets, GDP or public finances, unless these effects are fully justified.

On the contrary, **TAG2013** seems to be more prone to introduce indirect effects, but always considering that they are context specific; the type and magnitude of economic impacts which occur will depend upon the scheme type and more importantly the local attributes, such as workforce skills, the availability of land for development, etc.

A.3. Measurement and aggregation of benefits and costs

Most of the reviewed manuals and handbooks make a clear distinction between financial and economic evaluation. **CEDEX2010** states that economic evaluation is performed by comparing the social costs and benefits of a transportation project, once homogenized along time through their social net present value (NPV_S). On the other hand, the financial evaluation only compares income and monetary costs associated with the project, calculating their financial net present value (NPV_F).

According to **EU2015**, financial analysis should, as a general rule, be carried out from the point of view of the infrastructure owner, in constant (real) prices and using a financial discount rate and covering a period appropriate to the project's economically useful life and its likely long term impacts. If, in the provision of a general interest service, owner and operator are not the same entity, a consolidated financial analysis, which excludes the cash flows between the owner and the operator, should be carried out to assess the actual profitability of the investment, independent of the internal payments. Alternatively, the key concept in economic analysis is the use of shadow prices to reflect the social opportunity cost of goods and services, instead of prices observed in the market, which may be distorted and require the conversion from market to shadow prices, and evaluation of non-market impacts and correction for externalities.

Interestingly, in **TAG2013**, a complete set of model appraisal tables, requesting an extremely detailed analysis of monetised costs and benefits, is recommended. There are also tables for public accounts (the taxpayers' view) and others that monetize external effects. Recommendations about the presentation of results are less detailed in **ADB2013** and **EIB2013**, although in both cases – as it corresponds to multinational

financing institutions – the role of financial analysis is very relevant. The **EU2015** guidelines also pay great attention to the financial profitability and financial sustainability of the projects.

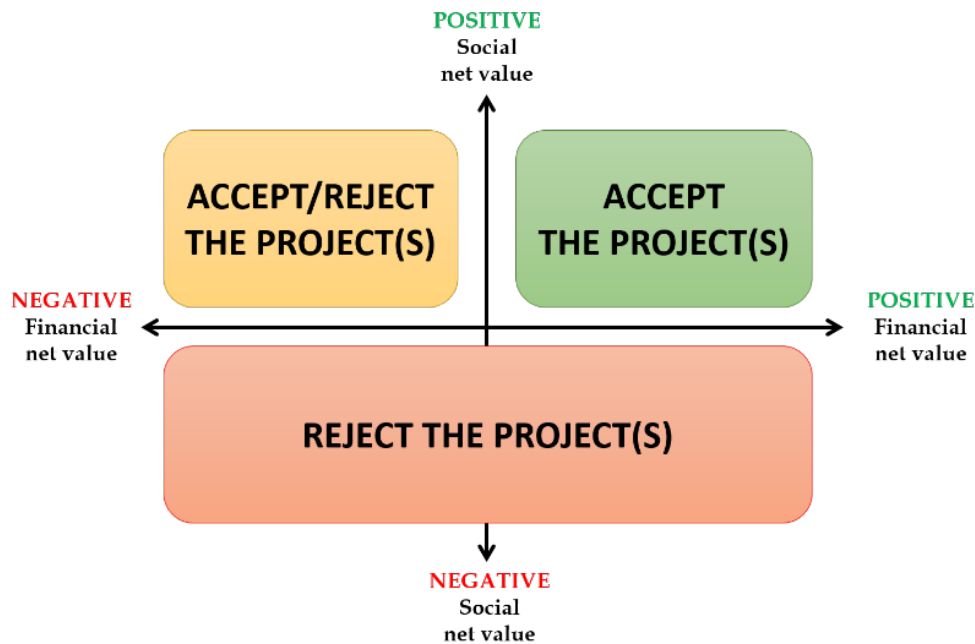
With regard to the definition of shadow prices and the economic valuation of the opportunity costs of the resources, **CEDEX2010** set out some general rules to correct distortions associated with taxes and lack of competition. For example, prices for input and output should be considered, generally, net of direct and indirect taxes. Despite these general rules, when indirect taxes (or subsidies) are intended as a correction for externalities it is justified to include them in project costs (benefits), provided that they adequately reflect the underlying marginal cost, but the appraisal should avoid double counting (e.g. including both energy taxes and estimates of full external environmental costs).

EU2015 suggests additional corrections to value project inputs at shadow prices. If they are tradable goods, border prices are used; if not, the alternatives include using standard conversion factors (according to trade differentials), applying *ad hoc* assumptions (depending on the specific hypotheses made on market conditions) or calculating a shadow wage (for labour). In the case of outputs, users' marginal WTP, remains as the best and more theoretically correct approach.

A.4. Decision criteria and the interpretation of results

Although it is a key issue in project appraisal, not all the reviewed manuals devote specific sections to clarify how the decisions should be made. In the case of **EU2015**, as showed in **Figure A.2**, the algorithm is quite simple: if $NPV_F > 0$ the project does not require financial support from the EU and – with some exceptions – no further assessment is performed. In the standard case where the project does require support ($NPV_F < 0$), the economic evaluation is carried out, and approved only if $NPV_S > 0$, which means that the society is better off with the project.

Figure A.3. Decision criteria under certainty conditions



Source: CEDEX2010

In **CEDEX2010** this decision procedure about accepting or rejecting a project is represented in **Figure A.3**, where financial and social values are both simultaneously considered. When the project is socially desirable (positive social *NPV*) but does not generate sufficient funds to attract private sector (negative financial *NPV*), the society must carry out the project only if there are no relevant budget restrictions by the government. If there were such restrictions, the society might have to reject this project. When the decision is to rank different projects, the criteria used should be based on the above principles. The society should prioritize projects with higher social values and always delay or reject those whose contribution to social welfare is lower, taking into account in any case the existence of budget constraints.

This is the implicit decision mechanism in most manuals. Only in **ADB2013** it is clearly criticized, arguing that it does not take into account the distributive effects of the project. In general, these guidelines prefer to complement the decision by using multi-criteria analysis (MCA), of a qualitative nature, rather than accepting only the results of the CBA. In **TAG 2013**, the primary metric used in reporting the CBA results in most circumstances is the benefit-cost ratio (BCR), which requires a clear definition of what constitutes the Present Value of Benefits (PVB) and Present Value of Costs (PVC). The general principle is that the PVC should only include impacts on the 'Broad Transport Budget', that is costs and revenues which directly affect the public budget available for transport. All other impacts, including operating costs and revenues for private sector transport providers and impacts on wider government finances, should be included in the PVB. In **EIB2013**, MCA is also admitted where full CBA, or other more standard quantitative appraisal techniques are not possible.

A.5. Other issues

There are many other specific topics that are covered to some extent in one or more of the reviewed guides. Possibly, the most relevant one relates to the treatment of risk, either with respect to the identification of its different sources or with regard to its inclusion in the calculation of the benefits and costs. While **CEDEX2010** considers that risk is inherent in the evaluation process and should be incorporated into it from the beginning (even in the decision criteria), the rest of the manuals prefer to deal with it as an additional element after the traditional CBA (without risk).

EIB2013 and **ADB2013** follow a very standard approach based on private investment analysis, implicitly using the capital-asset pricing model (CAPM), whereby the discount rate applied to the stream of future benefits and costs is adjusted by the risk premium corresponding to the expected volatility of such streams, volatility being taken as a measure of risk. In general, the type of risk analysis that can be applied to a given project depends on the quality of the data available to the analyst. If possible, a full risk analysis (with Monte Carlo simulations of the critical variables) is suggested. Otherwise, a scenario-based comparative may be useful, including the identification of switching values. Real options analysis is also admissible when uncertainty is particularly high and investments are irreversible.

EU2015 also considers that a risk analysis is mandatory to conduct the CBA of transport projects. The recommended steps for assessing the project risks include a sensitivity analysis, a qualitative risk analysis, and a probabilistic risk analysis, which should be completed with risk prevention and mitigation measures. The manual acknowledges that risk assessment is a complex function, but it should be always the basis for risk management, including how to allocate them to the parties involved and which risks to transfer to institutions such as insurance companies. In **TAG2013** detailed guidance is provided on how to deal with risk in forecasting values for the variables used in the project assessment and how to define the different values under which they must be compared and reported.

A second topic that is dealt with in some detail in some of the guides analysed relates to the choice of the discount rate, both social and financial, to compute the net present values. **ADB2013** presents a detailed survey on this issue, discussing its theoretical background and finding that there are significant variations in social discount rate policies practiced by countries around the world, with developing countries in general applying higher rates. ADB's policy on this item follows the World Bank approach and applies, with few exceptions, a single minimum rate of 10–12%. **EIB2013** argues that, as social time preferences might differ across countries, social discount rates might also differ, and in line with this view, for projects in the EU the EIB uses as a reference a real (that is, inflation-adjusted) social discount rate ranging from 3.5% to 5.5%, depending on the degree of maturity and expected growth rate of the national economy.

In **EU2015** two annexes are respectively devoted to discuss the main issues regarding financial and social discount rates. They acknowledge the existence of different methods and assumptions, but finally recommend that, in the first case, a 4% discount

rate in real terms is considered as the reference parameter for the real opportunity cost of private capital in the long term. Values differing from this benchmark may, however, be justified on the grounds of international macroeconomic trends and conjunctures, the Member State's specific macroeconomic conditions and the nature of the investor and/or the sector concerned. For social discount rates, 5% is used for major projects in Cohesion countries and 3% for the other Member States. Member States may establish a different benchmark if a justification is provided for this reference on the basis of an economic growth forecast and other parameters and its consistent application is ensured across similar projects in the same country, region or sector. In any case, the Commission encourages national governments to provide their own benchmarks in their guidance documents and to apply them consistently in project appraisal at national level.

Finally, there are other additional CBA topics that are covered to a different extent in the five guides.⁵¹ In the case of the value of time (VOT) and for the EU in particular, the already reviewed (see **Section 6**) *Handbook on the External Costs of Transport* only provides updated references only for road transport. **Table A.6** summarizes these values for Spain and the corresponding UE average.

Table A.6. Value of time references: Spain and UE average

Passenger transport	Urban		Interurban	
	Commuting - business	Personal	Commuting - business	Personal
Spain	12,1	5,6	15	5,6
UE Average	13,67	6,32	16,74	6,32
Freight transport	Value per tonne		Value per HGV/coach driver	
Spain	1,8		19,5	
UE Average	1,7		18,2	

Source: European Commission (2019). Values in €2016 (per person or per tonne).

⁵¹ Many of these elements are very example-specific and do not always follow the same criteria or definitions in their expositions. For these reasons, a case-by-case reading would be recommended instead.

ANNEX B. A THEORETICAL FOUNDATION OF THE COST-BENEFIT ANALYSIS RULES

Consider a representative consumer with a continuous and increasing utility function that only depends on the amounts chosen within a set of n consumption activities, including all goods and services produced in the economy,⁵² $U(x_1, \dots, x_n)$, where x_j represents the quantity of good or service j , with $j = 1, \dots, n$. This consumer chooses his optimal set of consumption activities by maximizing his utility, given his budget constraint. This constraint delimits all the combinations of goods and services that may be purchased at any given time, according to their (exogenous) market prices and consumer's income, which has two components.

On one hand, the consumer obtains income by working. Let us denote by \bar{l} the maximum time endowment, available for the consumer (for example, 24 hours per day, or 365 days per year), and by t_j the time required to consume each unit of good or service j . Denoting by w the wage received per unit of working time, consumer's labour income is given by wl , where l represents the working time chosen by the representative consumer, which is defined by the difference

$$l = \bar{l} - \sum_{j=1}^n t_j x_j.$$

On the other hand, we will assume that all firms are ultimately owned by this representative consumer and that they distribute all their profits; thus, the consumer's total income from profits is given by:

$$\Pi = \sum_{j=1}^n \pi_j,$$

where π_j is the maximum profit obtained by firm j from producing and selling good or service j . From each firm's point of view, this profit is obtained by solving the standard maximization program:

$$\max_{l_j} \pi_j = p_j x_j^s - w l_j = p_j f_j(l_j) - w l_j, \quad (\text{B.1})$$

where p_j is the market price of good or service j , and l_j represents the amount of labour (the only input in this model) used by firm j to produce x_j^s through the production function $f_j(l_j)$. Assuming that all the required equilibrium properties hold, the first order condition of this problem is given by:

⁵² The model developed in this annex provides a more formal derivation of all the expressions used in **Section 2**. It is based on Johansson (1993), Johansson and Kriström (2016), and de Rus and Johansson (2019).

$$\frac{\partial \pi_j}{\partial l_j} = p_j \frac{df_j(l_j^*)}{dl_j} - w = 0, \quad (\text{B.2})$$

and it allows us to obtain as a solution $\pi_j = p_j f_j(l_j^*) - w l_j^*$. Note that in this equilibrium the sum of all the labour inputs used by firms must be equal to the working time offered by the representative consumer, that is,

$$\sum_{j=1}^n l_j^* = l.$$

We can now use these results to finally define the consumer's budget constraint, which is given by:

$$\sum_{j=1}^n p_j x_j \leq \Pi + w l, \quad (\text{B.3})$$

which can be also rewritten as:

$$\sum_{j=1}^n p_j x_j \leq \Pi + w \left(\bar{l} - \sum_{j=1}^n t_j x_j \right),$$

that is:

$$\sum_{j=1}^n g_j x_j \leq \Pi + w \bar{l}, \quad (\text{B.4})$$

where $g_j = p_j + w t_j$ represents the generalized price of good or service j .

It can be noted that expressions **(B.3)** and **(B.4)** are equivalent and, thus, we can write consumer's budget constraint in terms of market prices, $p = (p_1, \dots, p_n)$, and consumers' income, $\Pi + w l$, or alternatively, in terms of the generalized prices, $g = (g_1, \dots, g_n)$, and the maximum income (profits income plus the value of time endowment), $\Pi + w \bar{l}$.

We are now prepared to solve the individual's decision problem. If the utility function satisfies the non-satiation property, the budget constraint is binding, and the consumer's maximization problem reduces to:

$$\begin{aligned} & \max_{x_1, \dots, x_n} U(x_1, \dots, x_n) \\ & \text{s.t.} \quad \sum_{j=1}^n p_j x_j = \Pi + w l, \end{aligned} \quad (\text{B.5})$$

or, equivalently, in terms of generalized prices:

$$\begin{aligned}
& \max_{x_1, \dots, x_n} U(x_1, \dots, x_n) \\
& \text{s.t.} \quad \sum_{j=1}^n g_j x_j = \Pi + w\bar{l}.
\end{aligned} \tag{B.6}$$

This is the preferred expression to deal with in the economic evaluation of transport projects, since most of them can be generally interpreted as changes in generalized prices (either due to changes in market prices or time). The corresponding Lagrange function used to solve problem (B.6) is then given by:

$$L = U(x_1, \dots, x_n) - \lambda \left(\sum_{j=1}^n g_j x_j - \Pi - w\bar{l} \right), \tag{B.7}$$

which can be also rewritten as:

$$L = U(x_1, \dots, x_n) - \lambda \left(\sum_{j=1}^n g_j x_j - \sum_{j=1}^n p_j f_j(l_j^*) - w \sum_{j=1}^n t_j x_j \right). \tag{B.8}$$

First order conditions are given by:

$$\begin{aligned}
\frac{\partial L}{\partial x_j} &= \frac{\partial U(x^*)}{\partial x_j} - \lambda(g_j - wt_j) = 0, \\
\frac{\partial L}{\partial \lambda} &= \sum_{j=1}^n g_j x_j^* - \Pi - w\bar{l} = 0,
\end{aligned} \tag{B.9}$$

with $j = 1, \dots, n$ and $x^* = (x_1^*, \dots, x_n^*)$.

The solution of the above maximization program yields the Marshallian demand function for each good or service j , given by $x_j^* = x_j(g, y^g)$, with $y^g = \Pi + w\bar{l}$.

By substituting all these demands in the (direct) utility function, we obtain the consumer's indirect utility function, defined as:

$$U(x_1^*, \dots, x_n^*) = V(g, y^g). \tag{B.10}$$

In addition, note that by replacing the Marshallian demands in the Lagrangian function and taking into account first order conditions, we have that, in equilibrium:

$$L^* = V(g, y^g) - \lambda \left(\sum_{j=1}^n g_j x_j^* - \Pi - w\bar{l} \right) = V(g, y^g), \tag{B.11}$$

and, therefore:

$$\frac{\partial L^*}{\partial y^g} = \lambda = V_y = \frac{\partial V(g, y^g)}{\partial y^g}, \tag{B.12}$$

showing that the Lagrange multiplier can be interpreted as the consumer's marginal utility of income.

Once the model has been solved and a consumption (and production) equilibrium has been found in each market j , how can we measure the effects of a transport project on social welfare? In our single representative consumer setup, the change in social welfare, dW , is just given by the change in the consumer's utility: $dW = dU$ and, thus, considering the direct utility function evaluated in the initial equilibrium, we can write:

$$dW = dU = \sum_{j=1}^n \frac{\partial U(x^*)}{\partial x_j} dx_j. \quad (\text{B.13})$$

Then, substituting the first order condition of the consumer's maximization program given by (B.9) into expression (B.13), we obtain:

$$\frac{dW}{V_y} = \sum_{j=1}^n (g_j - wt_j) dx_j = \sum_{j=1}^n p_j dx_j. \quad (\text{B.14})$$

Equivalently, if we use the indirect utility function, we get:

$$dW = dV = \sum_{j=1}^n \frac{\partial V}{\partial g_j} dg_j + V_y dy^g. \quad (\text{B.15})$$

First, notice that the partial derivative with respect to g_j of the Lagrange function given by expression (B.8) yields the following:

$$\frac{\partial L}{\partial g_j} = \sum_{i=1}^n \frac{\partial U}{\partial x_i} \frac{\partial x_i}{\partial g_j} - \lambda \left(x_j + \sum_{i=1}^n \frac{\partial x_i}{\partial g_j} g_i - \sum_{i=1}^n wt_i \frac{\partial x_i}{\partial g_j} \right),$$

and also that, in the optimum we have:

$$\frac{\partial V(g, y^g)}{\partial g_j} = \frac{\partial L^*}{\partial g_j} = \sum_{i=1}^n \left(\frac{\partial U}{\partial x_i} - \lambda (g_i - wt_i) \right) \frac{\partial x_i}{\partial g_j} - \lambda x_j = -\lambda x_j. \quad (\text{B.16})$$

Thus:

$$\frac{\partial V}{\partial g_j} = -\lambda x_j = -V_y x_j. \quad (\text{B.17})$$

which can be replaced into expression (B.15) to finally obtain a usable expression that allows to evaluate the effects of transport projects:

$$\frac{dW}{V_y} = -\sum_{j=1}^n x_j dg_j + dy^g. \quad (\text{B.18})$$

Now consider that the transport project, interpreted as a change in the generalized price of good or service j , is only due to a change in the market price p_j , while the required (travel) time t_j remains constant, that is, $dg_j = dp_j$. In this case we have:

$$dy^g = d(\Pi + w\bar{l}) = \sum_{j=1}^n \frac{\partial \pi_j}{\partial p_j} dp_j = \sum_{j=1}^n x_j^s dp_j. \quad (\text{B.19})$$

By substituting this result into expression (B.18), and assuming that all product markets clear, $x_j = x_j^s$:

$$\frac{dW}{V_y} = -\sum_{j=1}^n x_j dp_j + \sum_{j=1}^n x_j^s p_j = 0, \quad (\text{B.20})$$

that is, a change in the generalized price of good or service j due to a change in the market price p_j (with t_j constant) does not produce any effect on welfare.

Alternatively, consider now that the change in the generalized price of good or service j is due to a change in time t_j while the market price p_j remains constant, that is, $dg_j = wdt_j$. In this case:

$$dy^g = d(\Pi + w\bar{l}) = \sum_{j=1}^n w \frac{\partial \pi_j}{\partial t_j} dt_j = \sum_{j=1}^n w \left(p_j \frac{\partial f_j(l_j^*)}{\partial l_j} - w \right) \frac{\partial l_j}{\partial t_j} dt_j, \quad (\text{B.21})$$

which, according to the first order condition of the profit maximization program of firm j given by expression (B.2) is zero, i.e., $dy^g = 0$. Then, by substituting this into expression (B.18), we finally obtain that:

$$\frac{dW}{V_y} = -\sum_{j=1}^n x_j w dt_j, \quad (\text{B.22})$$

as discussed in detail in Section 2.

REFERENCES

- Ahlfeldt, G. M., Redding, S. J., Sturm, D. M. and Wolf, N. (2015):** “The economics of density: Evidence from the berlin wall”. *Econometrica*, 83 (6): 2127-2189.
- Becker, G. S. (1983):** “A theory of competition among pressure groups for political influence”. *The Quarterly Journal of Economics*, 98: 371-400.
- Brühlhart, M., and Mathys, N. A. (2008):** “Sectoral agglomeration economies in a panel of European regions”. *Regional Science and Urban Economics*, 38: 348-362.
- Carbo, J. M., Graham, D. J., Anupriya, Casas, D. and Melo, P. C. (2019):** “Evaluating the causal economic impacts of transport investments: Evidence from the Madrid-Barcelona high speed rail corridor”. *Journal of Applied Statistics* 46: 1714-1723.
- CE/INFRAS/ISI (2011):** *External costs of transport in Europe*. Delft: CE Delft.
- Ciccone, A. (2002):** “Agglomeration effects in Europe”. *European Economic Review*, 46: 213-227.
- Collier, P. and Venables, A. J. (2018):** “Who gets the urban surplus?”. *Journal of Economic Geography*, 18(3):523-38.
- Combes, P. P., Duranton, G. and Gobillon, L. (2008):** “Spatial wage disparities: sorting matters!”. *Journal of Urban Economics*, 63: 723-742.
- Combes, P. P., Duranton, G., Gobillon, L. and Roux, S. (2010):** “Estimating agglomeration economies with history, geology, and worker effects”. In: Glaeser, E. L. (ED.) *Agglomeration economics*. Chicago, IL: university of Chicago press.
- Combes, P. P., Duranton, G., Gobillon, L. and Roux, S. (2012):** “Sorting and local wage and skill distributions in France”. *Regional Science and Urban Economics* 42: 913-930.
- Christidis, P. and Brons, M. (2016):** *External costs of freight transport in European Union Member States*, s.l.: European Commission, Joint Research Centre (JRC).
- Dahlby, B. (2008):** *The marginal cost of public funds. Theory and applications*. The MIT Press.
- de Rus, G. (2010):** *Introduction to Cost-benefit Analysis: Looking for Reasonable Shortcuts*. Edward Elgar Publishing, Cheltenham, UK.
- de Rus, G. (2011):** “The BCA of HSR: Should the government invest in high speed rail infrastructure?”. *The Journal of Benefit-Cost Analysis*, 2(1):1-28.
- de Rus, G. (2015):** “La política de infraestructuras en España. Una reforma pendiente”. *Policy Papers* 2015-08, FEDEA.
- de Rus, G. and Johansson, P. O. (2019):** “Measuring the economic effects of transport improvements”. *Policy Papers* 2019/01, FEDEA.
- de Rus, G. and Socorro, M. P. (2010):** “Infrastructure investment and incentives with supranational funding”. *Transition Studies Review*, 17(3): 551-567.

- de Rus, G. and Socorro, M. P. (2014):** “Access pricing, infrastructure investment and intermodal competition”. *Transportation Research Part E*, 70: 374-387.
- de Rus, G. and Socorro, M. P. (2017):** “Planning, evaluation and financing of transport infrastructures: Rethinking the basics”. *Review of Network Economics*, 16(2):143-160.
- de Rus, G. and Socorro, M. P. (2019):** “Pricing and investment in alternative transport infrastructures”. *Transportation Research Part A*, 119: 90-107.
- DfT (2016). TAG UNIT A2.1:** Wider Economic Impacts Appraisal. London: HMSO.
- Di addario S. and Patacchini E. (2008):** “Wages and the city. Evidence from Italy”. *Labour Economics*, 15: 1040-1061.
- Dupuit, J. (1995):** “De la mesure de l'utilité des travaux publics (1844)”. *Revue Française d'Économie*, Programme National Persée, 10(2): 55-94.
- Downs, A. (1957):** *An economic analysis of democracy*. New York: Harper& Row.
- Duranton G. and Puga, D. (2004):** “Micro-foundations of urban agglomeration” economies. In: Henderson, J. V., Thisse, J. F. (eds.) *Handbook of Regional and Urban Economics*. Amsterdam: Elsevier.
- Engel, E., Fischer, R. and Galetovic, A. (2014):** *The Economics of Public-Private Partnerships*. Cambridge University Press.
- Eddington, R. (2006):** *The Eddington Transport Study. Main report: transport's role in sustaining the UK's productivity and competitiveness*. London: The Stationery Office.
- European Commission (2019):** *Handbook on the external costs of transport*, Brussels.
- Florio, M., Morretta, V. and Willak, W. (2018):** “Cost-Benefit Analysis and European Union Cohesion Policy: Economic Versus Financial Returns in Investment Project Appraisal”. *Journal of Benefit-Cost Analysis*, 9(1): 147-180.
- Flyvbjerg, B. (2014):** “What you should know about megaprojects and why: An overview”. *Project Management Journal*, 45(2): 6–19.
- Flyvbjerg, B., Ansar, A., Budzier, A., Buhl, S., Cantarelli, C., Garbuio, M., Glenting, C., Holm, M.S., Lovallo, D., Lunn, D., Molin, E., Rønne, A., Stewart, A. and Van Wee, B. (2018):** “Five things you should know about cost overrun”. *Transportation Research Part A*, 118: 174-190.
- Freeman, A. (1993):** *The measurement of environmental and resource values: theory and methods*, Resources for the Future, Washington D.C.
- Gahvari, F. (2006):** “On the marginal cost of public funds and the optimal provision of public goods”. *Journal of Public Economics*, 90:1251-1262.
- Graham, D. J. and Gibbons, S. (2019):** “Quantifying wider economic impacts of agglomeration for transport appraisal: existing evidence and future directions”. *Economic of Transportation*, 19: 1-21.

Graham, D. J., Gibbons, S., Martin, R. (2009): “Transport investments and the distance decay of agglomeration benefits”. *Working paper*, Imperial College of London.

Graham, D. J., McCoy, E. J. and Stephens, D. A. (2014): “Quantifying causal effects of road network capacity expansions on traffic volume and density via a mixed model propensity score estimator”. *Journal of American Statistical Association*, 109(508): 1440-1449.

Hanemann, W. M. (1982): “Applied welfare analysis with qualitative response models”, *CUDARE Working Papers*, 241. University of California, Berkeley.

Harberger, A.C. (1965): “Survey of literature on cost-benefit analysis for industrial project evaluation”. Paper prepared for the Inter-Regional Symposium in Industrial Project Evaluation, reprinted in A.C. Harberger, *Project Evaluation* (collected papers) (Midway Reprint Series), Chicago: The University of Chicago Press.

Harberger, A.C. (1976): *Project Evaluation (collected papers)*, Midway Reprint. University of Chicago Press. Chicago.

Holl, A. (2012): “Market potential and firm-level productivity in Spain”. *Journal of Economic Geography*, 12 (6): 1191-1215.

HEATCO (2006): *Developing Harmonised European Approaches for Transport Costing and Project Assessment* (HEATCO). Deliverable D5: Proposal for harmonised guidelines, Stuttgart: IER University of Stuttgart.

HM Treasury (2018): *The Green Book: Appraisal and Evaluation in Central Government*. London: Treasury Guidance.

INFRAS/IWW (2004): *External costs of transport: Update study*, Karlsruhe & Zürich: Infrass & IWW.

Johansson, P. O. (1993): *Cost-benefit analysis of environmental change*, Cambridge University Press, Cambridge, MA.

Johansson, P. O. and Kriström, B. (2016): *Cost-benefit analysis for project appraisal*. Cambridge University Press, Cambridge, MA.

Johansson, P. O. and Kriström, B. (2019): “On the Formulation of the Alternative Scenario in Cost-Benefit Analysis”, *CERE Working Papers* 2019:6. Centre for Environmental and Resource Economics. Umeå University and Swedish University of Agricultural Sciences.

Kahneman, D. and Tversky, A. (1979): “Prospect theory: An analysis of decisions under risk”. *Econometrica*, 47 (2): 263-291.

Laffont, J. J., Rey, P. and Tirole, J. (1998): “Network Competition: I. Overview and Nondiscriminatory Pricing”. *The RAND Journal of Economics*, 29 (1): 1-37.

Laffont, J. J. and Tirole, J. (1993): *A Theory of Incentives in Procurement and Regulation*, MIT Press, Cambridge. MA

Lancaster, K. J. (1966): “A new approach to consumer theory”. *Journal of Political Economy*, 74: 132-57.

- Li, H., Graham, D.J. and Majumdar, A. (2013):** “The impacts of speed cameras on road accidents: an application of propensity score matching methods”. *Accident Analysis and Prevention*, 60: 148-157.
- Mackie, P. J., Jara-Díaz, S. and Fowkes, A.S. (2001):** “The value of travel time savings in evaluation”. *Transportation Research Part E*, 37: 91–106.
- Mackie, P., Graham, D. J. and Laird, D. (2012):** “Direct and wider economic benefits in transport appraisal”. In Andre de Palma, Robin Lindsey, Emile Quinet, Roger Vickerman (eds), *Handbook in Transport Economics*. Edward Elgar, London, 501-526.
- Mackie, P. J., Worsley, T., and Eliasson, J. (2014):** “Transport appraisal revisited”. *Research in Transportation Economics*, 47: 3-18.
- Maibach, M., Schreyer, C., Sutter, D., van Essen, H. P., Boon, B. H., Smokers, R., Schroten, A., Doll, C., Pawlowska, B. and Bak, M. (2008):** “Internalisation Measures and Policies for All External Cost of Transport (IMPACT)”. *Handbook on estimation of external costs in the transport sector*, Version 1.1 Report Delft.
- Marrocu, E., Paci, R. and Usai, S. (2013):** “Productivity growth in the old and new Europe: The role of agglomeration externalities”. *Journal of Regional Science*, 53: 418-442.
- Martin, P., Mayer, T. and Mayneris, F. (2011):** “Spatial concentration and plant-level productivity in France”. *Journal of Urban Economics*, 69: 182-195.
- Melo, P. C., Graham, D. J. and Noland, R. B. (2009):** “A meta-analysis of estimates of urban agglomeration economies”. *Regional Science and Urban Economics*, 39: 332-342.
- Mion, G. and Naticchioni, P. (2005):** “Urbanization externalities, market potential and spatial sorting of skills and firms”. *Cepr discussion papers* 5172. Centre for economic performance, London School of Economics and Political Science.
- Mohring, H. (1972):** “Optimization and scale economies in urban bus transportation”. *American Economic Review*, 62(4): 591-604.
- Mohring, H. (1976):** *Transportation Economics*. Ballinger Publishing Company.
- Navrud, S., (2002):** *The state-of-the-art on economic valuation of noise*. Final Report to European Commission DG Environment.
- Neuberger, H. (1971):** “User benefit in the evaluation of transport and land use plans”. *Journal of Transport Economics and Policy*, 3: 52-75.
- Niskanen, W. A. (1971):** *Bureaucracy and representative government*. Chicago: Aldine-Atherton.
- Quinet, E. (2003):** “Short term adjustments in rail activity: The limited role of infrastructure charges”. *Transport Policy*, 10(1): 73-79.
- Robinson, J. A. and Torvik, R. (2005):** “White elephants”. *Journal of Public Economics*, 89: 197-210.

Savvides, S.C. (1994): “Risk analysis in investment appraisal”. *MPRA Paper* 10035, University Library of Munich, Germany, revised 14 Aug 2008.

Small, K.A. (1999): “Project evaluation”. In Gómez-Ibáñez, J.A., Tye, W.B. and W. Clifford (eds.): *Essays in Transportation Economics*. Brooking Institution Press, Washington DC.

Sobel, R. S. (1998): “The political costs of tax increases and expenditure reductions: Evidence from state legislative turnover”. *Public Choice*, 96: 61-80.

Stigler, G. (1971): “The theory of economic regulation”. *Bell Journal of Economics*, 2 (1): 3-21.

Venables, A.J. (2007): “Evaluating urban transport improvements: cost-benefit analysis in the presence of agglomeration and income taxation”. *Journal of Transport Economics and Policy*, 41: 173-188.

Viscusi, W.K. (2018): *Pricing lives. Guideposts for a safer society*. Princeton University Press, Oxford

Willig, R.D. (1976): “Consumer’s surplus without apology”. *American Economic Review*, 66(4): 589-97.

Zanakis, S. H., Solomon, A., Wishart, N. and Dubliss, S. (1998): “Multi-attribute decision making: A simulation comparison of select methods”. *European Journal of Operational Research*, 107: 507-529.