



PRO/CNMC/001/16 METHODOLOGICAL GUIDE ON THE EVALUATION OF STATE AID

June 9th, 2016

Index

I. Introduction	3
II. Definition, rationale and effects of State aid.....	6
II.1. Definition.....	6
II.2. The normative rationale for State aid.....	7
II.3. Other motives behind the existence of State aid.....	10
II.4. The effects of State aid.....	11
III. <i>Ex ante</i> evaluation of State aid	15
III.1. Necessity: market failures or objectives of common interest	16
III.2. Appropriateness and minimum restriction.....	17
III.3. Proportionality: Cost-Benefit Analysis.....	18
III.4. Other approaches to <i>ex ante</i> evaluation	20
IV. <i>Ex post</i> evaluation of State aid	23
IV.1. Formal aspects of evaluation.....	24
IV.2. Methodological aspects of evaluation	29
V. Concluding remarks.....	47
 Appendix 1. Market failures and State intervention	 50
 Selected references	 53

I. Introduction

State aid control is one of the distinctive elements of the European Union (EU) competition policy, in order to ensure aid regimes' compatibility with the Single Market. The system of control has undergone a process of modernisation¹ in recent years. This has implied a bigger decentralisation, i.e. greater flexibility for Member States in terms of *ex ante* notification while setting an enhanced *ex post* control (*monitoring*) through the evaluation of some schemes² with the highest potential impact on competition.

Regardless of this supranational prescription, evaluation is still essential and should be fostered at a national level too. The Spanish Constitution³ establishes that public resources allocation should fulfil the criteria of equity and efficiency. This principle has permeated government policies with a budgetary cost⁴.

Within those public policies, State aid is especially relevant, both in quantitative and in qualitative terms⁵. Despite the aim of addressing some objectives of common interest, State aid can alter market efficiency and general welfare when they are not necessary or proportionate.

The Spanish legal framework⁶ includes some of these aspects, especially the need for respecting the principles of good regulation and minimum competitive distortions.

¹ COM(2012) 209 : Communication on EU State Aid Modernisation

² Commission Regulation (EU) n°651/2014 of 17 June 2014 declaring certain categories of aid compatible with the internal market in application of Articles 107 and 108 of the Treaty (GBER, General Block Exemption Regulation) and Communications on sectoral Guidelines impose evaluation on schemes with an annual budget above €150 million (in some fields) and could advise evaluation as well in other programmes, which are novel, which affect sectors under technological or regulatory changes or which exhibit a high degree of sectoral or geographic selectivity.

³ Article 31 of 1978 Spanish Constitution.

⁴ Articles 26.2 and 70.2 of the Law 47/2003, of 26 November, of General Budgetary issues.

Articles 8 and Additional Provisions 13 and 18 of Law 38/2003, of 17 November, on Subsidies and Articles 10 and 15 of Royal Decree 887/2006, of 21 July, which passes the Regulation developing that Law.

⁵ The latest annual reports by CNMC on State Aid show a sizeable support to the financial sector and gradual decline in regular aid, from €5,553 million in 2009 (0.52% of GDP and close to €120 *per capita*) to €2,779 million in 2013 (0.28% of GDP and slightly below €60 *per capita*).

⁶ Article 8.2 of Law 38/2003, of 17 November, on Subsidies and Articles 10 and 15 of Royal Decree 887/2006, of 21 July, which passes the Regulation developing that Law. In the competition legal framework, Article 11 of Law 15/2007, of July 3, on the Defence of Competition, Article 7 of the Regulation on the Defence of Competition (approved by Royal Decree 261/2008, of 22 February)

de Defensa de la Competencia y arts. 7 y 8 del RD 261/2008 de 22 de febrero por el que se aprueba el Reglamento de la Ley señalada. También en esta misma línea, el art. 18 de la Ley 20/2013 de 9 de diciembre, de garantía de la unidad de mercado. En otros países de la UE, por

limits *ex post* assessment to formal aspects and does not take advantage of experiences from other countries⁷ which allow for a more strict control on State aid.

For instance, Articles 44 to 51 of Law 38/2003, of 17 November, of Subsidies set a framework for financial control which could lead to partial or total reimbursement of the aid if some irregularities are found. However, this is just a formal oversight which does not focus on efficiency and competition and which only applies to subsidies (leaving aside other forms of State aid, such as tax exemptions, loans or guarantees). The Law 47/2003, of 26 November, of General Budgetary issues, apart from developing the framework for financial control (in its Articles 140 to 175), does refer to the evaluation of expenditure policies (in its Articles 70 to 72) but only in a general manner.

The **National Commission for Markets and Competition (CNMC)**, as a **competition authority, is entitled to assess State aid effects on competition**. The compatibility of every aid regime with competition law is judged by the European Commission, but national authorities, like the CNMC, still have a crucial role in terms of advocacy. These duties are set in Article 5.1.e) of Law 3/2013, of 4 June, which creates the National Commission for Markets and Competition and Article 11 of Law 15/2007, of 3 July, on the Defence of Competition, being developed by Article 7 of the Regulation on the Defence of Competition (approved by Royal Decree 261/2008, of 22 February).

Therefore, the **CNMC is the appropriate institution to elaborate a guide with some principles for aid evaluation**. That is why the **CNMC Action Plan for 2015** (in its strategic action 8.5) set the commitment to “**draft a Guide to assess the effectiveness of state aid**”.

This guide seeks to target **three types of audience**. First, **public bodies and employees that are responsible**, from a technical point of view, **for managing State aid in their daily work**. The culture of evaluation has to permeate all the layers of public policy and, more specifically, State aid, shaping its design, implementation and monitoring.

Secondly, **public administrations that are responsible**, from a political point of view, **for approving State aid schemes and other public policies**. Evaluation is essential not only as an element of transparency and accountability but also as a learning tool about public policies, in the belief that ***what cannot be measured, cannot be improved***. The importance of evaluation is such that, in the fields of State aid and public policy, there are experts who deem that **intervention of the government should be conditional on having a robust evaluation**.

ejemplo en Dinamarca, la autoridad de competencia tiene competencias incluso para instar la eliminación de un régimen de ayudas públicas que dañe seriamente la competencia (OCDE, 2005).

⁷ For instance, in Denmark the competition authority may urge to remove a State aid regime which impairs competition (OECD, 2005).

Thirdly, **the public in general**, in order to increase awareness about the need for evaluating State aid and other policies. Citizens have the right to demand an accountable and efficient government. Policy-makers must explain why a given policy was adopted and whether or not it reached its objectives. In this regard, **this guide aims to join the growing institutional and academic consensus on the increasing importance of evaluation in the fields of State aid and public policy.**

Against this backdrop, this guide is organised as follows. After this first introductory section, the second section addresses the definition, justification and effects of State aid. Then the third section deals with *ex ante* evaluation and the fourth section with *ex post* evaluation. The fifth section summarizes the main concluding remarks.

II. Definition, rationale and effects of State aid

II.1. Definition

The **Treaty on the functioning of the European Union** (TFEU), in its Article 107(1), defines State aid as ‘any aid granted by a Member State or through **State resources in any form** whatsoever which distorts or threatens to **distort competition** by **favouring certain undertakings** or the production of certain goods in so far as it **affects trade between Member States**’. There are several elements central to this definition (European Commission, 2014a):

- **State resources granted in any form:** be it through grants, tax exemptions (or deferrals), equity participation, soft loans or guarantees. State aid could also take place in the sale or purchase of assets, goods and services when the most economically advantageous tender has not been selected due to the lack of an open, transparent, public and non-discriminatory process. State aid may also appear when granting exclusive rights or when compensating for the provision of services of general economic interest (SGEI), whenever the entrustment act has not followed a competitive pattern and remuneration is not in sync with market rates.
- A **selective economic advantage for one (or more) undertaking(s)**, an idea where there are several implicit factors:
 - Competition policy (State aid rules contained therein) generally only applies to an ‘undertaking’. This notion is absent in the TFEU but has been shaped by European Courts’ jurisprudence as ‘any entity engaged in an economic activity, regardless of its legal status and the way in which it is financed’. The concept of ‘economic activity’ only excludes the provision of goods and services on a non-economic basis, i.e. guided by principles of authority (involving the exercise of public powers) or solidarity (as in social security, health care, some education and research activities and infrastructure not meant to be commercially exploited).
 - The undertaking favoured by the aid enjoys an economic advantage it would not have obtained under normal market conditions in the absence of State intervention.
 - The measure grants, with some degree of selectivity, a specific advantage to certain undertakings or categories of undertakings or to the production of certain goods and services or economic sectors.
- The measure **distorts or threatens to distort fair competition** between undertakings, as the position of aid beneficiaries is strengthened.
- All these factors may **affect intra-EU trade** and other aspects of the EU Single Market (given the impact on other firms’ decisions like investment or establishment).

This last component of the definition is crucial to understand why State aid is perhaps the most idiosyncratic element of EU competition policy, with a minimal degree of decentralisation for the Member States.

But there are also other international *fora* engaged in this topic, chiefly the **World Trade Organization** (WTO). Its Agreement on Subsidies and Countervailing Measures (SCM) defines a ‘subsidy’ as a ‘financial contribution by a government or any public body which confers a benefit’. But the SCM only applies to subsidies which are ‘specific’, i.e. ‘a subsidy available only to an enterprise, industry, group of enterprises or group of industries’. Relying on these concepts, subsidies are divided into two categories (WTO, 1994):

- **Prohibited subsidies:** the specificity of the subsidy is conditional on export performance (“export subsidies”) or on the use of domestic over imported goods (“local content subsidies”). These two categories of subsidies are prohibited because they are designed to directly affect trade.
- **Actionable subsidies:** production subsidies whose specificity is at the level of an enterprise, an industry or a region. These categories of subsidies are not prohibited but can be challenged should they cause adverse effects to another WTO country.

The novelty of the WTO’s is that it allows a country to use the dispute settlement mechanism to seek the withdrawal of a subsidy or the removal of its adverse effects (including the ability to apply a ‘countervailing duty’ on subsidized imports).

II.2. The normative rationale for State aid

Generally, the allocation of resources should be left to the market, as it ensures that welfare and efficiency are maximized, from three standpoints:

- ✓ **Allocative efficiency:** resources flow wherever they are most valuable.
- ✓ **Productive efficiency:** firms use the most efficient technology.
- ✓ **Dynamic efficiency:** agents are incentivised to go on seeking efficiency, as those which do not use resources efficiently are moved out of the market.

Therefore, the “**raison d’être**” for State aid or for any other public intervention (such as taxation or regulation) is the **existence of ‘market failures’**, where the market is not necessarily the optimal mechanism to allocate resources. Market failures can be divided into four categories:

- **Externalities:** the consumption/production of a good/service entails some social costs/benefits, which go beyond the private costs/benefits for the consumer/producer. There are negative externalities, like environmental pollution or the systemic risk created by a single financial institution, and positive externalities, like training or research. In order to reach the optimal output, one option for the government would be to grant State aid for certain goods, services

or technologies, although there are other alternatives, such as regulation or taxation.

- **Public goods:** whose consumption is non-excludable (i.e. it is difficult to actually impede access to them) and non-rival (i.e. use made by an agent does not affect use by others). Public goods tend to be infra-provided by the market, so the solution is normally for the public sector to produce directly the good or to promote its private provision through subsidies or regulation. Nonetheless, public goods (like financial stability) are rare in purity and they tend to manifest in mixed forms, where the menu of options to correct market failures is bigger. On the one hand, there are club goods (excludable but non-rival), such as technology or not congested network infrastructures (in transport, telecommunications, energy or postal services). On the other, there are common goods (rival but non-excludable, and hence similar to externalities), like natural resources.
- **Imperfect information:** it is perhaps the most frequent market failure, as agents take usually their decisions without perfect knowledge. The case of asymmetric information is especially harmful for market efficiency. In those circumstances, there is room for state intervention, providing aid so that more transactions can be completed or regulating to promote an augmented flow of information (although the market itself can provide mechanisms to facilitate the exchange of information). Credit markets are a clear-cut sign of imperfect and asymmetric information, where the public sector grants subsidies, soft loans or guarantees so that more transactions can be completed.
- **Increasing returns:** they refer to the context where it is more efficient to concentrate production in a few firms (or even in one) because there are high fixed costs and average costs tend to decrease with the scale of production. The problem would be that, within this backdrop of market power, firms do not behave competitively, harming consumer welfare with higher prices and lower quantity. Public intervention should start firstly by tackling market power through liberalisation and competition policy. If those are not appropriate, the government could consider regulation or State aid to ensure an adequate supply at a reasonable price, as has been traditionally the case in network industries, like transport, telecommunications, energy and postal services.

Nonetheless, for all these public interventions to beat the market in terms of efficiency they have to be adequately designed, i.e. they have to target at the correction of the imperfection while minimizing additional or disproportionate distortions. Otherwise, the free play of market forces could be preferable as a second best, even if not leading to an optimal allocation.

Another motive for public intervention is improving the market outcome in terms of **redistribution**. Should some consumers (with low income or from specific regions) face difficulties to access certain goods or services, the government could use its resources (or its regulatory tools) to ensure an adequate supply at a reasonable price. Some examples, beyond the above-mentioned network industries, could be

regional development policy, aid after natural disasters or schemes of social support to individual consumers.

Actually, many of these public interventions are not affected by the State aid control legal framework (Friederiszick *et al*, 2007). For instance, transfers to individual consumers or compensations following natural disasters are compatible with the Single Market and are not appraised by the European Commission. In the same vein, public payments to remunerate the provision of Services of General Economic Interest (SGEI) are not labelled as State aid under certain conditions⁸. Finally, some activities (like public education or health care) are developed on the basis of solidarity and are not considered 'economic activities'. Notwithstanding that, these measures should also be assessed from the perspective of public policy evaluation, even if many of these exceed the scope of this Guide.

Finally, there are also **non-economic reasons** behind public intervention, be it through State aid or through other means. Within these purposes, we can find culture promotion or heritage conservation, although some of these may partially have an economic dimension (if they are considered as externalities).

⁸ Like an entrustment act following competitive procedures or a compensation calculated with the costs of an efficiently managed firm, without overcompensation (coherently with the Court of Justice ruling on the 'Altmark case' and its subsequent development by EU regulatory packages on SGEI).

Table 1. Reasons and alternatives for public intervention

Reason behind the intervention		Specific example	Alternatives for intervention	
			State aid	Other means
Market failures	Externalities	Training (+)	Subsidies, tax deductions, public provision	Certification to ensure quality
		R & D & i (+)	Subsidies, tax deductions, soft loans, guarantees	Patent system incentivizing innovation
		Environmental pollution (-)	Subsidies, tax discounts and soft loans to green investments	Taxes for polluting activities, regulation to limit certain activities, property rights
	Public goods⁹	Financial Stability	Recapitalisation, guarantees	Regulation to discourage adequate risk taking
		Technology	Subsidies, tax deductions, soft loans, guarantees	Patent system incentivizing innovation while ensuring access
	Imperfect information	Credit markets (credit crunch for SMEs, exports...)	soft loans, guarantees (for SMEs, exporting activities)	Information, initiatives to broaden capital markets and create alternative sources of finance
	Increasing returns	Network industries	Subsidies, soft loans to foster investment	Regulation to ensure adequate supply and affordable access
Other objectives of general interest	Redistribution	Regional development	Soft loans for investment	Supply-side reforms to foster development
	Non-economic goals	Culture & heritage	Subsidies	Sanctions for heritage deterioration

II.3. Other motives behind the existence of State aid

In practical terms, and from a positive point of view, State aid also appears owing to **less normative or ideal motives** (Neven and Röller, 2000). The most traditional purpose of public subsidies has been **industrial policy**, whereby governments try to expand output, employment and/or investment in specific areas, sectors or firms. Industrial policy can be more ‘offensive’ or rather ‘defensive’ (OECD, 2010). The former aims to promote the development of new activity, e.g. fostering business creation or tapping foreign markets. The latter tries to protect existing

⁹ Environment and (uncongested) network infrastructures are other examples of public goods, but they have been dealt with, respectively, in externalities and increasing returns. In turn, financial stability, herein considered within public goods, is also considered an externality.

activity, preventing firm closure, employment losses and sometimes even sheltering domestic firms from foreign competition.

Even if some of these goals would appear to be noble, we'll see in the next section that this policy is largely ineffective to achieve them (OECD, 2010), chiefly because of the risk of '**subsidy races**' (similar responses by other geographic areas). There are as well huge costs on efficiency, mainly owing to the high degree of selectivity.

Actually, the selectivity of these measures gives a hint of other not so noble reasons behind the existence of State aid, related to political economy issues and 'government failures'. The support to a specific activity or region can be the result of **lobbies**, **political pressures** or the **electoral cycle**. Again, in the coming section we will grasp benefits and costs of these measures, especially harmful when they do not respond to normative criteria, like efficiency or redistribution.

II.4. The effects of State aid

II.4.a. Benefits of State aid

Aid schemes that are well designed and targeted and that **respond to normative criteria**, like the **correction of market failures**, are **likely to generate net benefits for overall** welfare if they achieve a **more efficient allocation of resources**. Sometimes, State aid (or other forms of public intervention) might be the only way to improve market outcomes and attain optimal results (in terms of prices, quantity, quality or innovation) if there are externalities, public goods, imperfect information and increasing returns. Furthermore, some measures add to its direct effect (the change in incentives which triggers a different course of action by the beneficiary to address the market failure) some positive indirect effects, like spillovers in the case of R & D and crowding-in effects in investment-support schemes¹⁰.

If there is a **redistributive aim or non-economic factors** behind aid schemes, it is essential to try to quantify the benefits in economic terms¹¹, so that they can be compared to the costs entailed. Because, as we will describe later on, even those measures which respond to normative criteria and market failures imply potential drawbacks, justifying the need for a thorough evaluation.

And **when public policy is not driven by this normative intervention logic**, **benefits tend to vanish** while **shortcomings soar**. For instance, as far as industrial policy is concerned, if an aid scheme were to be successful increasing a sector's output in a given area, other regions would be tempted to replicate that scheme. Finally, multiple governments could initiate a 'subsidy race', largely ineffective in fostering economic activity, as no region would increase its relative

¹⁰ Schemes that promote subsidized investment, through regional development or through infrastructure (like broadband), may foster further deployment efforts by private firms.

¹¹ This assessment of distributional outcomes from an economic perspective may be challenging, if not impossible, and is likely to involve judgements of an ethical nature (Friederiszick *et al*, 2007).

competitiveness vis-à-vis the others (and with high costs to public finance, efficiency and competition). Needless to say, the issue of ‘subsidy races’ is of utmost importance within the EU Single Market and the Commission’s strict control of State aid is one tool to minimize the risk of this kind of behaviour among Member States.

That is why this use (and abuse) of sectoral/vertical industrial policy is deemed to be outdated, to the benefit of an horizontal perspective consisting in allowing the free play of competitive forces, so that successful firms and sectors contribute to output, employment and investment. A well-established result in theoretical and empirical economics is that the market is better than the government in ‘picking winners’.

And, finally, when State aid is the consequence of lobbies or political drivers (unconnected to efficiency), the balance is even bleaker. There are actual benefits from this type of measures, but they are very limited and they are concentrated in the very few agents who lobbied for the scheme and are not seeking the maximisation of total welfare (but theirs). Meanwhile, costs are enormous and they are spread over many agents, as the next subsection describes.

II.4.b. Costs to public finances

First and foremost, State aid implies a **substantial drain on budgetary resources** (OECD, 2010). Currently, most advanced economies face daunting challenges for their public finances, both in the short term (given the need to put fiscal aggregates on a sound footing) and in the long term (given an aging population and high levels of public debt). Hence, it is probably the right moment to be more careful when granting State aid by evaluating which schemes outperform others.

Even without those strains on public finances, State aid regimes must be very thoroughly assessed. Their setbacks go well beyond direct administrative costs, as there is an **opportunity cost of public resources** (which could be used for alternative purposes), which have to be financed through **tax hikes and/or expenditure cuts**¹² (OECD, 2010). On the one hand, raising taxes entails a direct administrative cost but more importantly an indirect efficiency burden, as agents’ incentives are distorted and economic activity is harmed (some taxes are also regressive). On the other, spending reductions may be detrimental as well if resources are diverted from more productive uses for efficiency or redistribution, such as research and development, infrastructure, education or health care.

This efficiency losses provoked by raising distortionary taxation and pruning productive expenditure are very difficult to gauge, although there are estimates (OECD, 2010). For instance, the ‘deadweight efficiency loss’ of taxation could range between 18% and 24% of the amount levied for the United States, and these overheads are likely to be bigger for countries at a lower stage of development. Meanwhile, the social return of investing in education in OECD countries could be

¹² If aid is financed through the issuance of government debt, tax hikes and/or expenditure cuts will unavoidably arrive later, with the additional cost of higher interest payments.

around 8.5%. The problem is that, when assessing the optimality of an aid scheme from a partial equilibrium perspective (focusing exclusively on the scheme), it is impossible to guess which measure(s) would have been implemented should the aid have not been approved.

Given that the topic of public finances goes well beyond the scope of this guide, the approach followed in the next sections suggests focusing on the evaluation of efficiency and competition. So, when the net impact is positive from the standpoint of efficiency and competition, these gains should still be compared with an estimation of total (direct and opportunity) costs to public finance.

Consequently, bearing in mind the negative effect on public finances, one should err on the side of prudence when carrying out a cost benefit analysis of an aid regime. In other words, even when the assessment of a measure is positive from the competition and the efficiency perspective, one should remain very cautious before endorsing the measure's implementation, as direct and indirect implications on public finance could be large.

The fiscal impact is perhaps the most visible drawback of State aid, especially when there is a need for budgetary consolidation. However, the effects on competition and efficiency are not less relevant, especially in the long run. Free competition is a key driver of long-term growth, which contributes to improve public finances by improving tax receipts without raising rates.

II.4.c. Costs to competition

State aid could impact competition very adversely in these events (OFT, 2007):

- ✗ The **free play of competitive forces is altered**, especially in the case of a sectoral industrial policy, discriminatory *per se*. The first and most evident asymmetry appears between beneficiaries and their competitors in the same sector and the same region. But distortions spill over into other geographic areas (because of trade interconnections) and into other sectors with horizontal or vertical links. Needless to say, the impact on competition is even more negative when inefficient beneficiaries increase their market share to the prejudice of non-beneficiaries who are more efficient.
- ✗ Beneficiaries may increase their **market power** (where market share can be a proxy), laying the groundwork for an abuse of that dominant position regardless of competitive pressures coming from rival firms or consumers.
- ✗ The aid generates **consolidation dynamics** in the industry to an extent where competitive pressures recede and there is a **higher risk of collusion**. These consolidation trends can surge within beneficiaries (for instance, when the aid was conditional to firm size and companies decided to merge in advance in order to benefit from the aid) or within their competitors (seeking to gain size to offset the disadvantage they face vis-à-vis the beneficiaries).

- ✗ **Entry barriers** are raised (for instance, through the investment in over-capacity), so that operating firms are sheltered from actual and potential competitors. Indeed, the mere existence of State aid can be considered a barrier to entry, as it impacts asymmetrically incumbent firms and potential entrants.

II.4.d. Costs to efficiency

State aid could impact efficiency negatively in these aspects (OFT, 2007):

- ✗ **Allocative efficiency** could be harmed, as the ratio of relative prices is altered. When public intervention responds to a market failure, allocative efficiency can actually increase if the selected tool is appropriate (incentivising a change in the beneficiaries' conduct that allows fixing the market failure) and proportional to the goal sought. However, there may be cases where, despite a genuine aim to correct a market failure, the measure leads to undesirable side-effects, distortions of market prices and suboptimal decisions. Changes in the beneficiaries' conduct affect decisions by non-beneficiaries, so beneficiaries may again adjust their behaviour, triggering second and third round effects (OFT, 2007). There are also efficiency deadweight losses when the scheme does not achieve the 'incentive effect' that causes the needed change of conduct in the beneficiary (European Commission, 2014b). Actually, State aid may crowd out other expenditures and investments by beneficiaries or by (perhaps more efficient) non-beneficiaries. Needless to say, there is higher cost on efficiency in the case of sectoral initiatives, defensive industrial policies or schemes responding to lobbies or political pressures.
- ✗ **Trade and investment flows are altered** as the most extreme consequence of the allocative efficiency impairment. Distortions of market prices and suboptimal decisions affect not only beneficiaries and their direct and potential competitors but also firms and sectors with vertical links. This ends up affecting companies' decisions of where to produce and/or sell, setting and explanation for the European Union interest in State aid.
- ✗ **Productive efficiency** is also negatively affected as State aid provides extra financial resources that reduce the incentives to minimize cost. This problem is known as the 'soft budget constraint' (OECD, 2010) and is especially harmful when the schemes are perceived as permanent, automatically renewable and defensive (hindering the *Darwinian* mechanism by which inefficient firms close down). Furthermore, if competition is lessened and firms hold market power (as has been already mentioned), the so-called 'x-inefficiency' can appear (the disincentive to minimize costs given the dearth of competitive pressures).
- ✗ **Dynamic efficiency** is impaired if firms start to devote substantial amount of money in "subsidy-shopping", searching (or lobbying) for State aid instead of investing in drivers of competitive advantage (like research, investment or quality). Again, the toll on dynamic efficiency is even bigger when policy makers decide upon different measures following considerations unconnected to welfare and efficiency (as in the case of lobbies or the electoral cycle).

III. *Ex ante* evaluation of State aid

Ex ante evaluation assesses the programme's features and attempts to predict its impact using data and information before its implementation.

Relying on the assessment of benefits and costs abovementioned, every scheme of State aid must be subject to a thorough *ex ante* evaluation. This evaluation can be framed in three steps¹³ (Friederiszick *et al*, 2007, European Commission, 2014b):

1. **Necessity**¹⁴: the aid scheme must respond to a market failure or to an objective of common interest, such as redistribution or the pursuit of other non-economic goals.
2. **Appropriateness and minimum restriction**¹⁵: the aid scheme should be the optimal measure to tackle the market failure or to attain the objective of common interest. It must be weighed up against other designs of the aid scheme and against other forms of public intervention. And the existence of alternatives within the market to improve its outcome in terms of welfare should not be disregarded.
3. **Proportionality**¹⁶: the benefits of the aid scheme must outweigh its potential costs to public finance, efficiency and competition.

This setting mirrors the so-called '**State aid balancing test**¹⁷', laid out in the "State Aid action plan: Less and better targeted State aid: a roadmap for State aid reform 2005 – 2009" (European Commission, 2005, and CNC, 2008). Even if it is

¹³ This set of principles is contained in Recital 5 of the GBER, Commission Regulation (EU) n°651/2014 of 17 June 2014 declaring certain categories of aid compatible with the internal market in application of Articles 107 and 108 of the Treaty.

¹⁴ In EU law, this principle is included in Article 9.1(b) of the Directive 2006/123/EC, of the European Parliament and of the Council, of 12 December 2006, on services in the internal market. In Spanish law, it is alluded to in Article 129.2 of the Law 39/2015, of 1 October, of Common Administrative Procedure of Public Administrations. In those cases the application is limited to regulation, but this principle should inform every public sector intervention in the economy.

¹⁵ In EU law, although limited to the scope of regulation, this principle is implicit in Article 9.1(c) of the Directive 2006/123/EC, of the European Parliament and of the Council, of 12 December 2006, on services in the internal market. In Spanish law, limited as well to regulation, it is contained in Articles 129.2 and 129.6 of the Law 39/2015, of 1 October, of Common Administrative Procedure of Public Administrations. Specifically referring to State aid, it is alluded to in Article 8.2 of the Law 38/2003, of 17 November, of Subsidies.

¹⁶ In EU law, this principle is implicit in many Articles of the Directive 2006/123/EC, of the European Parliament and of the Council, of 12 December 2006, on services in the internal market. In Spanish law, it is alluded to in 129.3 of the Law 39/2015, of 1 October, of Common Administrative Procedure of Public Administrations. In those cases the application is limited to regulation, but this principle should inform every public sector intervention in the economy.

¹⁷ In Spanish law, Article 7 of the Regulation on the Defence of Competition (approved by Royal Decree 261/2008, of 22 February) states that the National Competition Authority must carry out a State aid balancing test when assessing the appropriateness of a given instrument, its 'incentive effect, or its necessity and proportionality.

considered as a single test, its complexity advocates taking those three steps sequentially (Friederiszick *et al*, 2007).

For instance, if a measure did not pass the first test, there would not be the need to check neither its appropriateness nor its proportionality. As has been already explained, when a measure does not fall within the normative rationale for public intervention, its drawbacks will certainly exceed its limited (if any) paybacks.

In the same vein, once a measure has fulfilled the first trial, it is logic to place the appropriateness test before the proportionality (cost-benefit) analysis. Once the most suitable measure has been found, it is the moment to ponder whether benefits still more than offset the costs of this public intervention.

III.1. Necessity: market failures or objectives of common interest

The **existence of a market failure or the attainment of an objective of general interest** is a *sine qua non* to even consider public intervention through State aid or through any other form. This elemental rule is essential to ensure that state funds (and more broadly public policies) are managed in a way that increases welfare and efficiency.

Therefore, every aid scheme has to identify and specify which is the market failure tackled or the objective of common interest pursued. This is the best strategy to minimize the risk of schemes driven by not ideal motives, e.g. the lobby of a group to benefit from a sectoral industrial policy.

Nonetheless, this strategy is not exempt from caveats. Chief among these is the fact that market failures are hard (if not impossible) to measure in economic terms (Friederiszick *et al*, 2007). And when the objectives of public intervention are redistributive or non-economic, the issue of measurement is even more challenging.

When economic concepts (such as efficiency) interact with other notions which go well beyond economics (such as equity), public decisions usually involve trade-offs, unless there are win-win situations (positive from an efficiency and an equity standpoint). That is why it is advisable to express some dimensions of non-economic issues in economic terms, if possible¹⁸.

Finally, market failures are very common and are present in many economic transactions. Therefore, from a cost-benefit perspective, it is rationale to concentrate efforts on those which are more harmful to general welfare.

¹⁸ As was said before, an apparently non-economic goal as culture or heritage conservation can be framed as an externality problem.

III.2. Appropriateness and minimum restriction

Once the objective pursued is clear, **the measure has to be the most appropriate to target it**. In turn, this judgement must be organized in three steps which can be considered sequential as well.

Firstly, the **aid scheme has to be the most suitable tool** to address that market failure or objective of common interest. This appraisal should be carried out **considering all the other feasible alternatives** to achieve these goals, be it through **public intervention** or through a more **market-oriented solution**. An evident application of this principle would be the replacement of sectoral subsidies' schemes by a horizontal industrial policy (supporting infrastructure, training or research) to foster employment in a region, or even consider broader economic reforms (such as supply-side reforms in labour and goods markets).

Secondly, the aid has to generate an **'incentive effect'**¹⁹. In other words, the aid itself should make the beneficiaries take a different course of action from the one they would have followed without the existence of the aid (European Commission, 2014b). This change in the beneficiaries' conduct is crucial to achieve the alleged benefits, like a higher efficiency through the correction of a market failure or another objective of common interest. A clear 'incentive effect' is again a *sine qua non* to consider the implementation of a State aid scheme, because otherwise the benefits of the measure would be low and outweighed by its costs (to public finance and to efficiency and competition).

An absence of an 'incentive effect' is a sign of 'deadweight losses of efficiency', as the government is subsidizing a conduct that the firm would have carried out anyway and public resources should be devoted to alternative uses where they are more productive. That is why the market failure argument should not be overstated when subsidizing an action which was already commercially profitable for the beneficiary (such as training, research or energy efficiency).

In the same vein, the assessment of the 'incentive effect' is essential as well to prevent 'subsidy races'. For instance, if a firm had already planned to develop a project and receives aid to base it in a given region, then the scheme has not an 'incentive effect' (the project was going to be carried out anyway and the aid just determined where), while certainly distorts efficiency, trade and competition (Claici, 2012).

Thirdly, we have to mull over the **optimal design of the aid**, i.e. it has to be proved that **those benefits** and the **'incentive effect' cannot be achieved with less amount of aid or through other less distortive scheme**. An example of this test could be the substitution of grants by more market-oriented financing (like soft loans or guarantees) in a research and development programme.

¹⁹ The existence of an 'incentive effect' is required by Article 6 of the GBER.

As we have seen, this last step involves to some extent a cost-benefit comparison exercise, so it could actually be integrated into the next phase.

III.3. Proportionality: Cost-Benefit Analysis

If the aid regime has passed the previous tests (there is an objective of efficiency or common interest, State aid is the most suitable way to address it, the aid scheme has an 'incentive effect' and its design is optimal), it is likely to maximize its benefits. Yet, these **benefits must be weighed against its potential costs**, which, *ex ante*, will depend on several factors (Friederiszick *et al*, 2007, OFT, 2007). These factors can in turn be divided into two categories: factors intrinsic to the aid and factors related to the market affected²⁰.

III.3.a. Factors intrinsic to the aid itself

- a.1. **Aid element** (or amount): if the aid entails costs (to public finance, to efficiency or to competition) it is obvious that these will be greater the bigger it is the 'aid element'. The concept of 'aid element' is more precise than the 'amount', as the former will be different depending on the instrument used²¹. Anyhow, the scheme's size has to be set in context of the market's magnitude, the time span, the firms' costs, the aid intensity (the part of the project that can benefit from State aid relative to the part that has to be financed by the firm itself) and other factors that are subsequently analysed.
- a.2. **Form of provision:** *a priori*, grants are more distortive than other instruments like guarantees or soft loans, whose 'aid element' is smaller and whose design is more compatible with the market. Nonetheless, whenever the scheme is not appropriate, the former statement would not hold. For instance, if there is moral hazard and the firms lack the right incentives, the guarantees would be called or the soft loans would not be recovered, triggering greater drawbacks.
- a.3. **Impact on the cost structure:** traditionally, operational aid to variable costs was deemed to be more distortive owing to its impact on price competition in the short-term. However, aid to investment or fixed costs may take a heavier toll on competition in the long run as it affects entry and exit decisions and it may create obstacles to potential competitors. So there is not a rule-of-thumb to ascertain which type of aid is more harmful, but it is easier to find a market failure (and therefore a potential justification for aid) in investment than in variable costs.

²⁰ In any case, the interpretation of some intrinsic factors (such as the costs affected by the aid) is interconnected with the extrinsic ones (characteristics of the market affected such as the existence of entry barriers).

²¹ For instance, in a grant scheme, the 'aid element' coincides with the 'used amount', but this is not the case for other instruments. According to the European Commission methodology, the 'aid element' for equity participation, soft loans and guarantees is calculated taking into account the profit for the beneficiary relative to market conditions.

- a.4. **Selectivity:** the more discriminatory is a measure, the more acute distortions it will provoke. This selectivity can be inherent to the scheme, like in sectoral industrial subsidies. Or worse, it could be due to the lack of a transparent open tender procedure because of bad management of public resources.
- a.5. **Other aspects of design:** the existence of schemes that are implemented once but that imply recurring payments for a relatively long time-span is very distortive. Firms that enter the market after the aid implementation have to compete with incumbent firms which are taking advantage of public subsidies. The problem may become even more acute when the aid regime aims precisely to prevent inefficient firms' closure.

III.3.b. Factors related to the market(s) affected:

- b.1. **Concentration and asymmetry of market shares:** the fewer the number of firms in the market and the higher the asymmetry between big and small enterprises, the worse it is likely to be the impact on competition. Again, this must be evaluated against other factors, especially the degree of discrimination introduced by the scheme (particularly, whether eligibility is conditional on firm's size, favouring or harming SMEs). In order to implement this principle, the trickiest issue is defining the relevant product and geographic market²².
- b.2. **Barriers to entry:** in sectors with sizeable obstacles to firm's entry and exit, State aid is likely to have a greater impact on competition. Barriers can be created endogenously by incumbents (through investment in overcapacity, R & D or brand-image), they can be inherent to the sector (like sizeable fixed costs in technological sectors) or they can be the consequence of public decisions (in the form of licenses or administrative procedures). Indeed, the mere existence of State aid can be considered a barrier itself, as it benefits incumbents while prejudicing small, nascent and potential competitors²³.
- b.3. **Product differentiation:** *a priori*, the higher the degree of differentiation, the lower impact the State aid is bound to have, as price competition is not so stiff. Nonetheless, in some markets with significant differentiation, State aid could hamper competition substantially if beneficiaries hold some market power and take advantage of the aid to build endogenous barriers to entry.

²² The first best would be taking advantage of competition policy cases or market research by marketing or consultancy firms. A second best would be using public statistics for a given sectoral branch (like the NACE classification) or accounting data on turnover by beneficiaries and their competitors (OFT, 2007).

²³ The prejudice to potential competitors is obvious: as they are not yet established they cannot take advantage of this public support and, at the moment of entering the market, they will have to compete with incumbent firms which have a long record of receiving aid (which in fact can be used to build endogenous barriers to entry, e.g. through investment in overcapacity). The prejudice to small and nascent firms happens when they are excluded from the scheme, be it directly (if aid is conditional on size or experience) or indirectly (given that applying for aid entails some administrative costs which weigh relatively more for small firms).

- b.4. **Other characteristics of the market:** the higher the extent to which trade or investment decisions are influenced by the aid, the bigger the potential to hamper competition and efficiency. Therefore, aid is more distortive in goods and services which are tradable and more prone to delocalisation. Other aspects to be born in mind could be the R & D intensity (connected to other factors such as differentiation or entry) and the integration with input markets (which could reveal the extent to which the distortions could spill over). And, finally, it should be assessed if the sector has been prone to a closer scrutiny by competition or even whether the aid regime itself (if already existed) was related to competition cases in the past.

III.4. Other approaches to *ex ante* evaluation

The approach followed for *ex ante* evaluation has only focused on its qualitative dimension hitherto. Nonetheless, there are also **tools to predict a programme's quantitative impact**. These exercises of quantitative *ex ante* evaluation may be warranted for programmes with a relevant size (for a given country or a region) or for whole lines of policy (such as financial support, regional development, R & D spending, broadband deployment, etc.) in order to estimate impacts on aggregate variables (such as GDP, employment or productivity) at a national or regional level.

On the one hand, when data or academic research are available, **statistic and econometric approaches** can be applied to estimate the impact of a given measure on aggregated variables in a bunch of countries/regions (see box 3.1), extrapolating then the results to our specific scheme. This exercise comes with some caveats, as including all the factors that affect the final outcome is challenging, extrapolating automatically the results might not be robust and the 'incentive effect' is not measured (the actual change in the firms' conduct caused by the public intervention itself). Yet these approaches still provide useful (if rough) benchmarks to foresee the likely effects of a given policy.

On the other hand, this *ex ante* quantitative evaluation can be done through **structural models**²⁴ (see box 3.2), which rely on restrictive assumptions about the behaviour of agents affected by the scheme and about the general functioning of the economy, especially identifying interactions between agents in different markets. Therefore, one must be cautious when analysing quantitative results arising from these estimations rather than taking them as an exact figure.

Nonetheless, these quantitative structural approaches **do constitute good complements to the other alternative tools for evaluation**. Firstly, they are valid robustness checks for both *ex ante* qualitative evaluation and *ex post* evaluation. Secondly, they address some shortcomings of *ex post* impact evaluation, as they provide a macroeconomic estimate (in contrast with the microeconomic approach at the level of the firm) and give hints of the channels through which State aid

²⁴ The most popular are DSGE (Dynamic Stochastic General Equilibrium) macroeconomic models, although there are also micro-simulation techniques.

effects take place²⁵. Finally, quantitative *ex ante* evaluations can also provide a framework for understanding how the aid scheme might play in a different economic environment²⁶ (Khandker *et al*, 2010).

Box 3.1. Quantitative evaluation through econometric approaches

Financial sector:

Koetter and Noth (2015). The results in this work estimate that banks which were likely to be beneficiaries of bail-out funds were able to expand slightly their interest margins (by increasing loan rates and depressing deposit rates), showing a reduction of competitive pressures. However, this widening of mark-ups was relatively small and did not hold a substantial impact on credit and deposit volumes, with market shares largely unaffected.

Regional development:

Mohl and Hagen (2010). Using data for the period 1995-2006, the authors find that EU regional policy only had a statistically significant and positive impact in 'Objective 1' zones (those whose income was below 75% of the EU average income). In those areas a 1% increase in EU funds led to an increase of GDP *per capita* by 0.1%-1.34%, depending on the specification.

Broadband deployment:

Qiang *et al* (2009). The authors find that, other things equal²⁷, an economy with 10 more broadband subscribers per 100 people enjoyed a higher average growth rate during the period 1980-2006, 1.21 percentage points (pp) for advanced countries and 1.38 pp for developing countries²⁸. Therefore the effect of broadband appears to be stronger than other telecommunication technologies, like fixed, mobile and conventional internet connections.

Czernich *et al* (2009). In a similar vein, this piece of research finds that for OECD countries, all else equal²⁹, a 10 pp increase in broadband penetration raises annual *per capita* growth by 0.92-1.45 pp. There is even a one-off effect on GDP *per capita* of 2.7-3.9 pp in the first year of broadband adoption.

Research and development:

Guellec and van Pottelsberghe de la Potterie (2001). The authors find an elasticity of 0.171 long-term GDP with respect to public R & D (slightly above business R & D, of 0.132, but well below foreign R & D, of 0.459) thanks to higher total factor productivity (TFP). Public R & D is especially productive in the upper education sector as opposed to defence expenditure³⁰.

²⁵ Structural models take into account interrelationships among endogenous variables, while impact evaluation is a reduced-form estimation to the extent that it only considers a one-way relation between exogenous variables (like programme's implementation and other factors) and endogenous ones (the programme's effects). This potential shortcoming of impact evaluation is not so worrisome in State aid, as it is under the discretionary control of the government and may be tagged as an exogenous variable. However, there might be cases where there are two-way relationships, e.g. a higher spending in a R & D aid scheme may generate higher income, which may lead to a new increase in R & D public funds.

²⁶ For instance, how the programme might work in a region or a context with lower/higher unemployment, lower/higher R & D spending, etc.

²⁷ They control for other fundamentals, such as investment, human capital and country dummies.

²⁸ Although the effect was not so statistically significant for developing countries.

²⁹ Again, the effect of other fundamentals (such as investment, human capital and country and time effects) is filtered out.

³⁰ Anyway, these estimates are very sensitive to the interaction between public and private funding. That is why it is crucial to appropriately measure the 'incentive effect', i.e. to what extent public support to R & D crowds in or crowds out additional private investment.

Box 3.2. Quantitative evaluation through structural models

Financial sector:

Kollman et al (2012). Given the sizeable support to the banking system, the authors try to gauge its effectiveness. A subsidy to the banking system equivalent to 1% of GDP (financed through lump-sum taxes³¹) could boost output by 1.17 percentage points (pp) in the first year with respect to the baseline scenario and by 0.29 pp in the fourth year. This is achieved through lower lending rates applied by banks, boosting investment (precisely the most subdued macroeconomic aggregate in financial crises). Therefore, banking support can be more effective than traditional fiscal *stimulus* through higher spending, as the latter tends to push interest rates up, crowding out private investment.

Roeger and in't Veld (2012). This exercise aims at estimating the relative performance of alternative options to support the banking system (again equivalent to 1% of GDP, but in this case financed through labour distortionary taxation) amidst a context of financial stress. The purchase of toxic assets reduces the risk premium and fosters investment, boosting output by 0.3 pp in the first year with respect to the no-intervention scenario, although this effect dissipates in the long run. Recapitalisation has similar channels but somewhat more positive effects, as it increases output by 0.5 pp in the first year and by 0.2 pp in the long-term with respect to the no-intervention scenario, with a lower fiscal cost given the remuneration earned by the government in financial markets (through dividends and/or capital gains). The role of guarantees is also deemed effective to prevent a collapse of GDP in the event of a default, but the long run effect is still negative given the need to increase distortionary taxes to absorb fiscal deficits.

Regional development:

Varga and in't Veld (2011). The authors find a positive long run effect of spending on cohesion policy, with a EU GDP 0.3 pp higher in comparison with a counterfactual scenario without this policy, even after accounting for the contractive impact of taxes to fund this higher expenditure³². Obviously the gains are higher for the largest net-recipient countries, although they also experience some setbacks in the form of higher inflation and external deficits. Meanwhile, net-contributor States suffer welfare losses, as they have to raise taxes to fund projects they do not take advantage of (even after taking into account that they benefit from spillovers owing to partners' growth). The outcome is also sensitive to the component of cohesion policy, where infrastructure investment is the fastest to bear fruits, while R & D and human capital investment take longer to materialize.

Research and development:

Roeger, Varga and in't Veld (2009). This work finds an impact of 0.3 pp in the long-term GDP of tax credit for R & D investment equivalent to 0.1% of the GDP (financed through higher lump-sum taxes³³).

³¹ This is an important caveat on the results, as lump-sum taxes do not capture the distortionary effects of taxation. Taking the latter into account, government support could be less expansionary or even contractive in the long run.

³² The authors assume that the policy is financed through labour taxation in an amount equivalent to 0.2% of each country's GDP.

³³ Again, this is an important caveat on the results, as lump-sum taxes do not capture distortionary effects of taxation.

IV. *Ex post* evaluation of State aid

Ex post evaluation is a systematic and objective assessment of a programme's impact once it is already being implemented or has actually finalized. *Ex post* evaluation is a comprehensive exercise that involves different tasks depending on the degree of ambition (Khandker *et al*, 2010):

- **Monitoring** involves identifying the programme's goals and setting key quantitative final indicators related to them, with targets to be achieved at specific points in time. Final indicators can be complemented by intermediate indicators, which are connected to them but at the same time are available more timely and respond more quickly to public interventions. Monitoring includes setting a full-blown system to track the programme's implementation (inputs) and performance (outputs) by comparing variables *ex ante* and *ex post* (*reflexive comparisons*), so it is a first step to promote accountability and engagement by all relevant stakeholders and to spot potential shortcomings in order to improve policy. Nonetheless, monitoring does not provide guidance on the actual effectiveness or efficiency of the programme.
- **Operational evaluation** aims to compare the final results and the measure implementation with what was planned *ex ante*. Therefore, it goes one step further than monitoring both in accountability and in policy-learning by finding potential gaps between what was intended and what was realized.
- **Impact evaluation** tries to measure the precise impact of the public policy, comparing the observed outcome after the programme's implementation with an unobservable counterfactual of what would have happened without State intervention. The cornerstone of this exercise is to measure (if and) how beneficiaries are affected by the policy, filtering out other effects coming from external factors. The '(self-)selection bias' must be addressed, i.e. the fact that beneficiaries already displayed differences from non-beneficiaries before the aid regime, those differences actually behind programme's participation. The aim of measuring causality (without drawing conclusions from mere correlation) makes this scope the most ambitious both in accountability and in policy improvement, given that the goal is to estimate the scheme's actual effectiveness. By measuring the intervention's impact and looking into the channels behind its effectiveness (or lack thereof), policy-makers and all the relevant stakeholders can assess whether the programme should be continued, tweaked, revamped or actually left aside.

These approaches are complements rather than substitutes (Khandker *et al*, 2010). For instance, without a good monitoring system to track indicators of the scheme's success, it is impossible to carry out a satisfactory operational and impact evaluation.

Therefore the *ex post* evaluation exercise, well beyond verifying the fulfilment of assumptions and the legality of the scheme³⁴, has to provide a **convincing and robust answer to the following questions**:

- Was the State **aid scheme appropriate** to achieve its objective of common interest? More specifically, did it succeed in creating an '**incentive effect**', generating a different course of action in the beneficiary, which was essential to attain the intervention's goals?
- Beyond those **direct effects** of the measure, were there '**indirect effects**', predicted or unpredicted, positive (such as spill-over effects in R & D schemes) or negative (such as distortions to efficiency, competition or trade flows)?
- Was the **measure proportionate**? In other words, did the scheme's benefits outweigh its costs? Besides, were the objectives achieved with the minimum restrictions? Or would it have been possible to reach the same target with a lower amount of aid through an alternative (less distortive) form of intervention?

The rationale for State aid *ex post* evaluation is manifold (European Commission, 2014b). Firstly, it serves the purpose of **transparency and accountability**. But mainly, it is a **learning** phenomenon to **improve the efficiency of public policies and resources**, assessing if, how and why a given measure works.

But for State aid *ex post* evaluation to be a fruitful exercise, some **formal and methodological patterns** ought to be followed.

IV.1. Formal aspects of evaluation

Beyond the quality of the methodological approach (which will be analysed afterwards), evaluation has to meet some standards of relevance, transparency and timeliness. This section divides these formal aspects into two issues: the schemes subject to evaluation and the content of the evaluation plan itself.

IV.1.a. Aid schemes to be evaluated

Wide-ranging evaluation, including impact assessment, is a time-consuming task which requires specialized human and technical resources. Therefore, from a cost-benefit perspective, this exercise should be **proportionate** and **focus on a manageable number of aid regimes**, according to several factors (Khandker *et al*, 2010, and European Commission, 2014b):

- **Size**: The bigger the amount, the higher the potential impact (on public finance, efficiency, competition and trade flows) and hence the more imperative the need for evaluation in order to improve general economic efficiency through an optimal design of State aid. This is the spirit of the GBER, whose Article 1.2³⁵

³⁴ As set in Recital 8 of the GBER.

³⁵ See also its Recital 8.

urges the elaboration of an Evaluation Plan for schemes whose annual budget exceeds €150 million, although this only applies to some categories³⁶. Other pieces of EU legislation³⁷ also mention the need for evaluation in certain schemes which grant a large amount of aid, so Decisions of the European Commission regarding compatibility of the regime can be made conditional to evaluation³⁸. Needless to say, national (and subnational) authorities could (and should) go further and carry out evaluation plans whenever the aid regime is sizeable relative to the granting authority's budget or vis-à-vis economic variables of the areas or sectors affected (even if the scheme is below the €150 million threshold set by the GBER to carry out a formal evaluation plan). Furthermore, the GBER only applies to some type of regimes (those exempted from notification to the European Commission) and to some categories, leaving still many schemes free from evaluation requirements, increasing again the need for higher ambition by national authorities.

- **Strategic importance:** the EU framework³⁹ includes other factors (beyond size) that may lead to evaluation, like novelty of the scheme, changes in the market affected (in its technology or regulation) or the degree of sectoral and geographic selectivity⁴⁰. Again, national (and subnational) authorities could (and should) go further and carry out evaluation plans under other circumstances, for instance when a regime is renewable, recurrent, permanent or likely to continue in the foreseeable future, as evaluation will be critical to improve it thenceforth.
- **Interest of the evaluation:** there are some schemes which, despite not being so relevant, provide an optimal groundwork for evaluation owing to the availability of high quality data. In those cases, from a cost-benefit analysis standpoint, evaluation might as well be implemented because of its low cost.

³⁶ Regional aid (except operating aid), aid for SMEs, aid for access to finance for SMEs, aid for R & D & i, aid for environmental protection (excluding reductions in environmental taxes) and aid for deployment of broadband infrastructures. Hence the remaining areas of the GBER are not covered by evaluation requirements: training, aid for disadvantaged workers, aid following natural disasters, aid for transport of residents of remote regions, culture and heritage conservation, sport and multifunctional recreational infrastructures and aid for local infrastructures.

³⁷ Like Guidelines for broadband deployment (Section 2.5 of [COM\(2013\) 25](#)), regional State aid (Section 4 of [COM\(2013\) 209](#)), risk finance investments (Section 4 of [COM \(2014\) 19](#)), aviation (Section 8.4 of [COM\(2014\) 99](#)), R & D & i (Section 5 of [COM\(2014\) 198](#)), environment and energy (Section 4 of [COM\(2014\) 200](#)) and rescue and restructuring (Section 6.7 of [COM\(2014\) 249](#)).

³⁸ Previously to the GBER, a few European Commission Decisions had already included the commitment of carrying out an evaluation of the scheme by the Member State. For instance, that was the case of Broadband Delivery UK framework scheme ([SA.33671](#)), with a total value estimated around €1.8 billion.

³⁹ Not the GBER but some of the horizontal and sectoral Guidelines applying to broadband deployment, regional State aid, risk finance investments, aviation, R & D & i, environment and energy and rescue and restructuring.

⁴⁰ The Guidelines for risk finance investments are very specific regarding the need to evaluate schemes with a very narrow focus on specific sectors and/or geographic areas (Section 4 of [COM \(2014\) 19](#)).

The conclusions drawn from this exercise could be subsequently applied to other more relevant schemes.

IV.1.b. The evaluation plan

The EU framework introduces interesting proposals which can be applied to any evaluation plan, regardless of its notification to the European Commission. The GBER sets the requirement to carry out the evaluation exercise relying on a plan approved by the Commission. Its article 2.16⁴¹ defines the 'evaluation plan'⁴² as a document containing at least the following elements (European Commission, 2014d):

1. **Objectives of the aid scheme to be evaluated.** The 'evaluation plan' has to specify which market failure or objective of common interest is addressed by the measure. The design of the regime should be commensurate with those objectives in terms of amount, duration, the instrument selected, aid intensity (and eligible costs), the beneficiaries targeted, the eligibility criteria and scoring rules for selecting them (such as size, sector, location) and the expected number of beneficiaries (and possible caps in the amount per beneficiary). Against this backdrop, the evaluation plan must state its expected impact, not only regarding the achievement of those general objectives but also at the level of the beneficiary (estimating the 'incentive effect' of the measure and also potential implications for non-beneficiaries).
2. **Evaluation questions.** The 'evaluation plan' ought to seek convincing and robust evidence on direct and indirect impacts and on proportionality and appropriateness.
 - **Direct impact:** the 'evaluation plan' has to include questions about the existence of the 'incentive effect', i.e. whether and to what extent the beneficiary has taken a different course of action from the one that would have happened without aid⁴³. In order to measure that, beneficiaries (and ideally non-beneficiaries) should be enquired about their plans and situation both before and after the introduction of the aid scheme. The evaluation exercise should also assess whether beneficiaries have been affected in a different fashion by the aid. Finally, the evolution of beneficiaries and non-beneficiaries relative competitive position (with indicators like market shares

⁴¹ See as well its Recital 8. Generally, the evaluation plan should be notified to the Commission within 20 working days following the entry into force of the scheme.

⁴² In Spain, in Article 8.1 of the Law 38/2003, of 17 November, of Subsidies, obliges granting authorities to prepare an strategic plan of subsidies with its aimed objectives, the terms needed to attain them, foreseeable costs and sources of financing. Nonetheless, this is a general plan instead of a programme-specific document.

⁴³ This inquiry about the 'incentive effect' is crucial because it can be measured more easily and robustly than indirect impacts. Furthermore, if there is not an 'incentive effect', it would be very difficult to justify that the measure has pay-offs (the change in the beneficiary's action, if any, is not caused by the aid, so there are windfall profits for beneficiaries) while it will certainly have costs to public finance, efficiency and competition.

or risk *premia*) may be tracked too, given that they also serve the purpose of measuring the indirect effects (such as those on competition).

- Indirect impact: the 'evaluation plan' must contain questions (for beneficiaries and ideally non-beneficiaries) about indirect effects, be them positive (like R & D spillovers) or negative (like crowding out activity and investment from other firms and regions). Relying on that information, the 'evaluation plan' must reach a conclusion on whether and to what extent were there effects on efficiency, competition and trade.
- Appropriateness and proportionality: the 'evaluation plan' should determine whether and to what extent the aid had the expected effects and reached its goals, at the level of the beneficiary and at the aggregate level of the objective of common interest sought. The proportionality involves considering if the benefits of the measure outweigh its drawbacks and whether the same goals could have been achieved with less aid or another (less distortive) instrument.

3. **Result indicators.** The answers to the previous questions must be provided through some indicators, preferably quantitative, objective and comparable. There will be some indicators (especially those measuring the impact on efficiency, competition and trade) that may be tracked in most types of schemes: change in competitiveness indicators of beneficiaries and non-beneficiaries (market shares, risk *premia*, etc.), the bias towards incumbents and big firms, trends in market power and concentration, variation in trade/investment flows, the risk of sectoral bias, etc. But in most cases, indicators will depend on the aim of the regime to be evaluated, as suggested in these illustrative examples (European Commission, 2014b):

- Regional aid: employment, investment and other economic variables in the beneficiary firms and regions, paying heed to the public funds leverage over additional private expenditure. In those schemes, it is of utmost importance including indicators of potential diversion of trade flows within regions or crowding-out of investment by other firms.
- R & D & i aid: R & D expenditure and other indicators such as patents, number of researchers and the leverage effect of public expenditure over private investment. These indicators should be tracked not only for beneficiaries but also for non-beneficiaries and at an aggregate level, in order to estimate potential spillover or crowding out effects.
- Environmental and energy aid: depending on the specific objective, the indicators could include the share of renewable energy, the mitigation of gas emissions, the introduction of standards or the increase in energy efficiency. However, beyond that direct impact, other potential effects should be followed, as environmental and energy aid has sizeable implications on firms and markets using these inputs. For instance, the electricity market could become more isolated from international competition, the introduction of renewable energy objectives could increase market power of some

generation companies and the introduction of environmental standards can affect trade flows.

- **Broadband aid:** the deployment of broadband in terms of population covered, connection speed, investment costs per household, leverage effect over private investment, effective demand by households, etc. As with aid to energy, impact indicators should also monitor vertically related markets, especially market power and concentration in telecommunication services in the affected regions.
 - **Aviation and airlines:** beyond direct impacts on economic activity (value added, passengers, flights, etc.), competition between airports should be assessed, in terms of the efficiency in using the infrastructure (to avoid subsidy races between airports that lead to duplication and over-capacity). The impact on market power and concentration in airlines should be appraised as well.
 - **Risk finance:** the leverage over private investment, the returns achieved by public and private funds and other measures of outreach (valued added, employment and firms affected). The impact on competition merits a special assessment in those cases, to check whether the scheme has benefited mostly nascent, small, innovative and high-productivity firms, which can exert competitive pressures on incumbent and big firms. The risks of sectoral or regional bias should also be tracked with objective measures.
 - **Rescue and restructuring:** beyond the effects on economic activity (firms, value added and employment affected), it is again essential heeding efficiency and competition indicators: changes in market shares, firms' birth and death rates (in order to see whether aid is preventing entry or exit), sectoral or other types of bias.
4. **Methodological issues.** The methodology envisaged must serve the purpose of measuring the actual impact of the State aid scheme. To that end, the evaluation must build a 'counterfactual' scenario of what would have happened in the absence of aid. These issues will be discussed separately in the next subsection given their complexity. Anyway, at this point it should be stressed that, within the methodological options available, the choice of the technique (sometimes determined by the availability of data) is not as important as other aspects of the evaluation, such as the selection of performance indicators (which has been dealt with above) or the collection of data (discussed below).
5. **Data collection.** Quantitative, objective, verifiable and comparable data are needed to answer the evaluation questions relying on the result indicators. If possible, it is preferable to use official sources (firm-level data from administrative records or statistical offices, subject to compliance with confidentiality obligations) instead of resorting to qualitative information or subject to manipulation (such as the one provided by questionnaires). Sometimes public information on companies can be available directly (like balance sheets and other economic and financial data) or through consultancy

or market research. Furthermore, it should be recalled that, in order to assess the actual impact of the scheme, both beneficiaries and non-beneficiaries data are needed. Hence, the obligation to provide information could be included within the eligibility criteria to be awarded the aid.

6. **Tentative timeline for evaluation.** *Ex post* evaluation should obviously start time after the entry into force of the measure, in order to get sufficient data which provide robust evidence on the aid scheme. But at the same time this must be balanced against the need to draw conclusions the sooner the better (Khandker *et al*, 2010), either to apply them to new regimes or to improve the scheme if still in place (publishing it and sending it to the European Commission by a certain date⁴⁴).
7. **Body to conduct the evaluation.** It is critical to ensure independence (in order to safeguard its impartiality vis-à-vis the granting authority) and competence (with adequate and appropriate human and economic resources)⁴⁵.
8. **Publicity.** Disclosure and outreach must be maximized, so that the granting authority and other relevant bodies are fully aware of the results. Thus, evaluation contributes to a better learning experience in order to improve State aid and public policy in general. Another relevant issue is that the data used to get the results should be made public (while respecting confidentiality issues) so that the evaluation can be replicated and criticized.

IV.2. Methodological aspects of evaluation

A robust and convincing *ex post* impact evaluation requires a **sound methodological approach**. In contrast to *ex ante* qualitative and quantitative assessment, *ex post* impact evaluation aims to measure the actual effect of a given State aid regime.

This is done within a **microeconomic scope**, i.e. **at the level of the beneficiary**, because this is the only way to appropriately **measure the ‘incentive effect’**, by **comparing the change in the beneficiary’s conduct with what would have happened without the aid**. Besides, it is interesting to go beyond the effects of the aid itself by looking into differences in impact depending on the amount of aid and on firms’ characteristics. Furthermore, it is important to recall that a microeconomic approach is not an impediment to reach conclusions on macroeconomic or aggregate impacts (on variables such as GDP, employment, productivity, R & D or even consumers’ welfare) at a national or regional scale, especially in programmes of a significant amount.

⁴⁴ For instance, the decision on the framework for broadband deployment in the UK ([SA.33671](#)) included the explicit obligation to release the evaluation results before the program expired, so that a hypothetical renewal could be assessed.

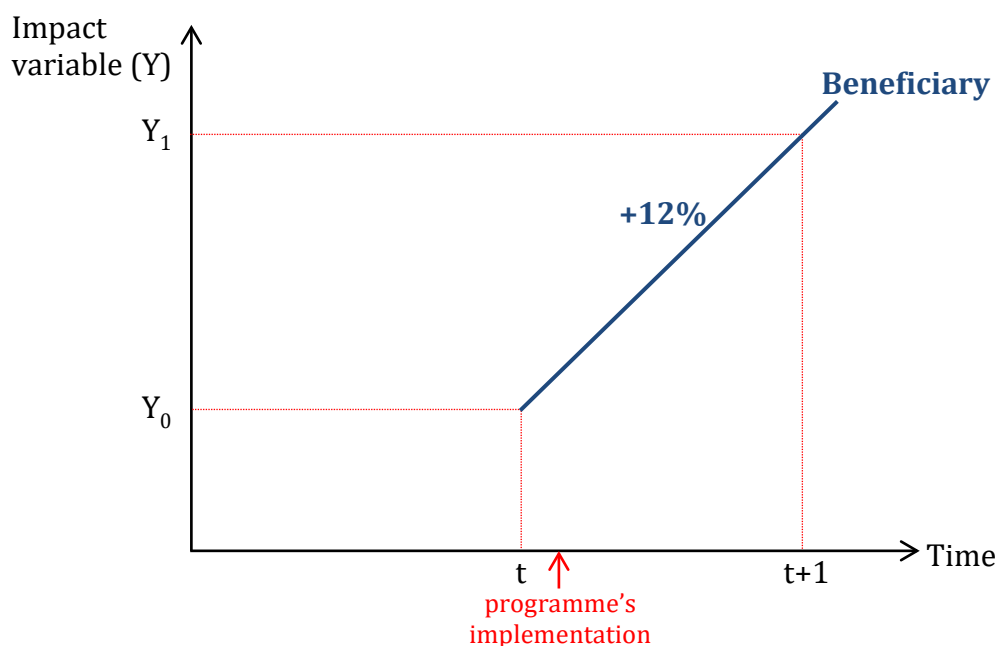
⁴⁵ The European Commission (2014d) considers that the Evaluation Plan must include the selection criteria even if the body has not been chosen yet.

As the outcomes following the aid implementation are observable, **the main challenge** is then finding and building the ‘**counterfactual**’ scenario in the **absence of aid**. This is crucial to disentangle variations in key variables that are actually triggered by the aid from those due to other factors. For that purpose, the evaluation exercise normally consists in comparing beneficiaries with a ‘control group’ of similar firms but which did not receive the aid, so that the difference between the two types of agents can be an estimate of the aid effect.

Before shedding light on how this impact evaluation can actually be carried out, we are going to explain **some potential mistakes in defining the ‘control group’ or the ‘counterfactual’**.

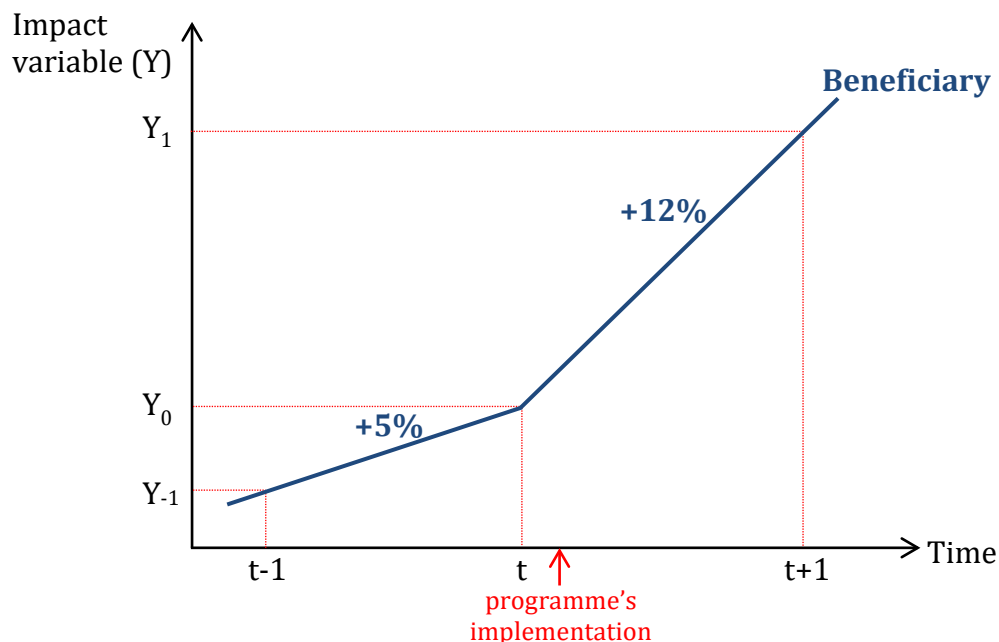
The first case would be represented by figure 4.1.a), where we **compare a firm’s outcome before and after receiving the aid**. Let us assume that the impact variable to be tracked (e.g. a firm’s R & D spending, investment or employment) rises at a 12% rate (from Y_0 to Y_1) after receiving the aid (between “t” and “t+1”). Interpreting that the aid has been effective (by boosting the variable’s growth by 12%, from Y_0 to Y_1) would be very probably an error of assessment.

Figure 4.1.a) Choosing a (likely) flawed counterfactual (past events)



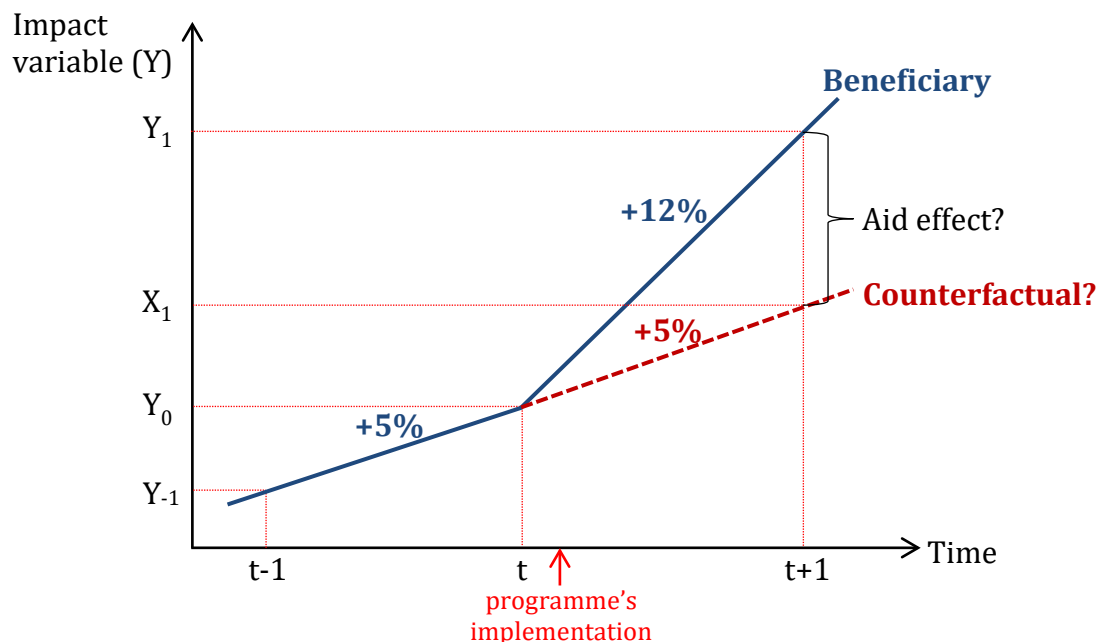
Let us assume, in figure 4.1.b), that the impact variable was already growing at an annual rate of 5% (from Y_{-1} to Y_0) before the programme’s implementation (between “t-1” and “t”). So the 12% increase (from Y_0 to Y_1) should not be exclusively attributed to the aid itself.

Figure 4.1.b) Choosing a (likely) flawed counterfactual (past events)



And even interpreting that the difference between the 5% *ex ante* growth and the 12% *ex post* is wholly due to the aid could be misleading, as figure 4.1.c) helps to understand. That reasoning would be equivalent to consider that what would have happened without the scheme (the 'counterfactual') between "t" and "t+1" is what actually took place between "t-1" and "t". In other words, we are assuming that without the aid the variable would have been X_1 at "t+1" (a 5% percent increase), so the effect of the aid ($Y_1 - X_1$) would be the difference between the actual value (Y_1) and the 'counterfactual' (X_1 , so that $X_1 = Y_0 + [Y_0 - Y_{-1}]$), an extra 7pp in the growth rate (from 5% to 12%).

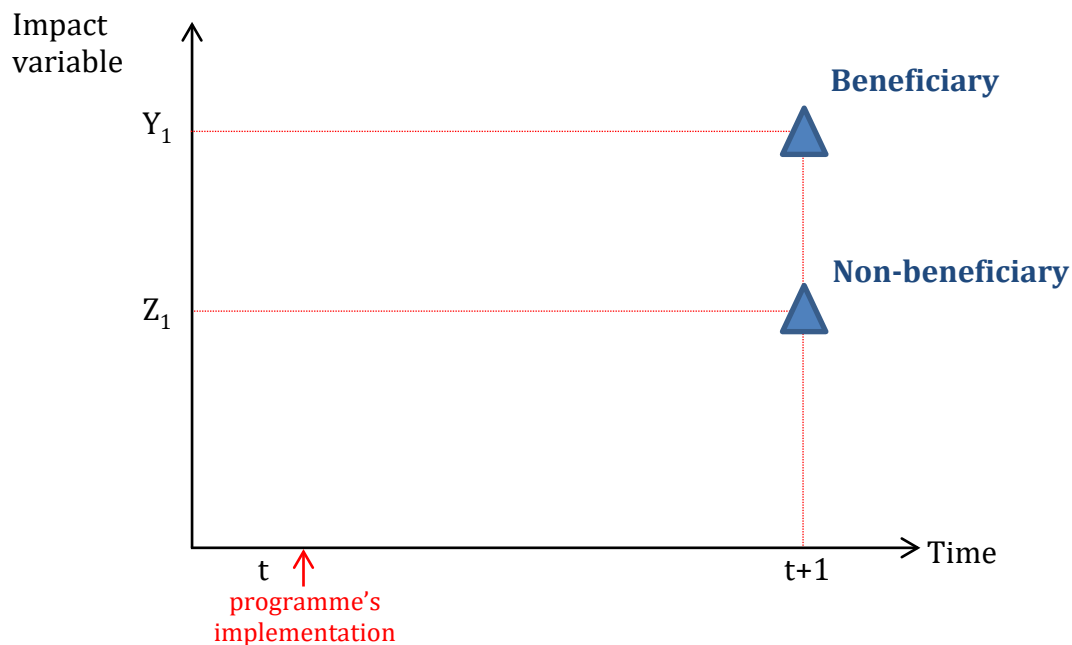
Figure 4.1.c) Choosing a (likely) flawed counterfactual (past events)



But this is not necessarily true, as the future rarely mirrors the past. The higher growth rate between “ t ” and “ $t+1$ ” could be due to an improved general environment or to an idiosyncratic shock which only affected the beneficiary (like a new and better management) but not necessarily related to the aid itself. Therefore, nothing can be concluded about a programme’s effectiveness relying on this type of analysis (it could well be the case that the scheme does not affect the firm at all or that the impact is even negative).

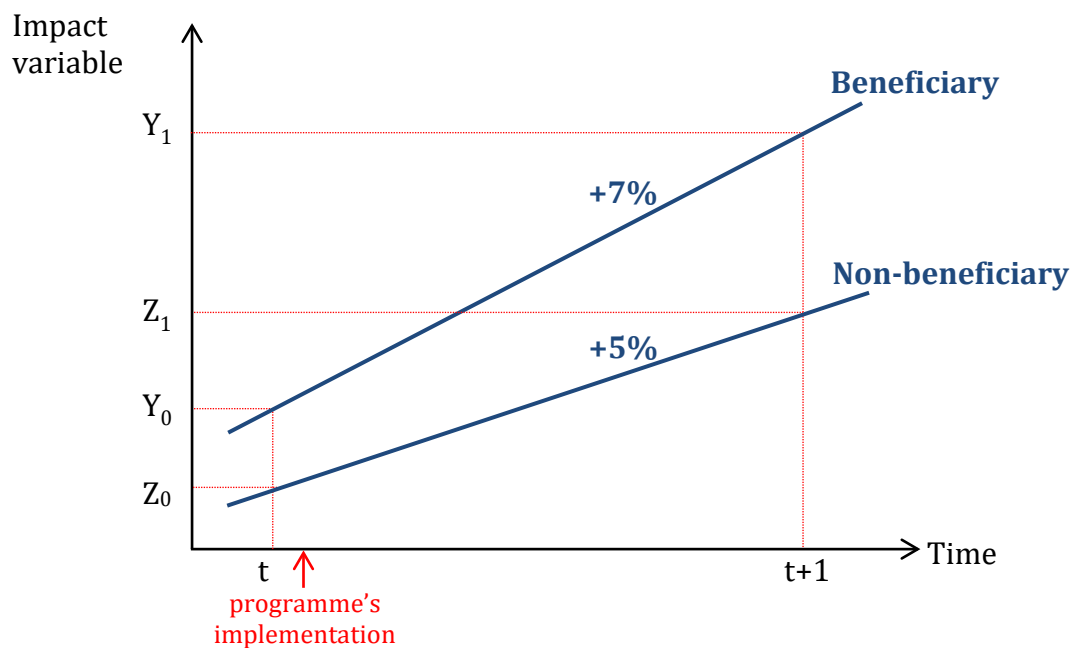
Other example of a wrong counterfactual, depicted in figure 4.2.a), would be a **direct and straightforward comparison between beneficiaries and non-beneficiaries** after receiving the aid. Let us assume that, after receiving the aid (at “ $t+1$ ”), the outcome variable is at Y_1 for a beneficiary and at Z_1 for a non-beneficiary. Again, one should not interpret that the aid is necessarily effective (by moving the outcome variable from Z_1 to Y_1).

Figure 4.2.a) Choosing a (likely) flawed counterfactual (non-beneficiary)



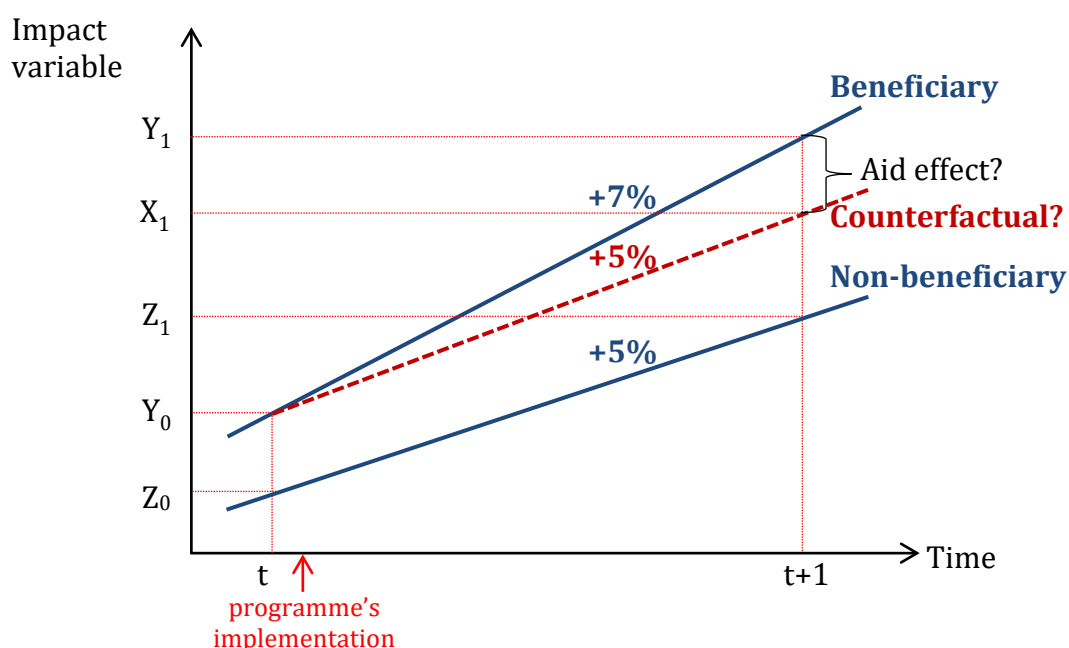
Let us assume, in figure 4.2.b), that beneficiaries and non-beneficiaries already departed from different values (Y_0 and Z_0 respectively) before the programme's implementation (at "t"). So the difference in outcomes after the aid (from Z_1 to Y_1) should not be exclusively attributed to the aid itself.

Figure 4.2.b) Choosing a (likely) flawed counterfactual (non-beneficiary)



And even interpreting that the aid the only factor behind the difference between the 5% growth for non-beneficiaries and the 7% for beneficiaries could be wrong. As figure 4.2.c) shows, that reasoning would be equivalent to consider that what would have happened to the beneficiary without the scheme (the counterfactual) is what has actually happened for the non-beneficiary. In other words, we are assuming that without the aid the variable would have been X_1 at “t+1” (a 5% percent increment), so the effect of the aid ($Y_1 - X_1$) is the difference between the actual value (Y_1) and the ‘counterfactual’ (X_1 , so that $X_1 = Y_0 + [Z_1 - Z_0]$), an extra 2pp in the growth rate (from 5% to 7%).

Figure 4.2.c) Choosing a (likely) flawed counterfactual (non-beneficiary)



But again this would not provide a convincing evidence on the programme’s effectiveness or lack thereof, as there are many differences between the beneficiary and the non-beneficiary beyond the fact of being granted of State aid. Higher growth for the beneficiary maybe existed well before the programme’s implementation (in “t-1” and before). Or, even if firms had a similar performance in the past, the difference between “t” and “t+1” could be due to a specific shock which only affected the beneficiary but unrelated to the aid scheme itself.

This is **especially relevant for State aid** because of the so-called ‘**selection bias**’. The *ex ante* differences between beneficiaries and non-beneficiaries actually explain programme’s participation, be it exogenously (because eligibility is conditional upon firm’s characteristics like size or region) or endogenously (because a firm which already has more productive projects and better management *ex ante* is more likely to apply for a programme and be awarded funds). Therefore, it will be challenging to disentangle *ex post* dissimilarities (owing to the aid itself) from *ex ante* ones (which already existed before the programme’s implementation). The latter could be confused with the former as beneficiaries tend to behave differently from non-beneficiaries and it is tempting to assign it

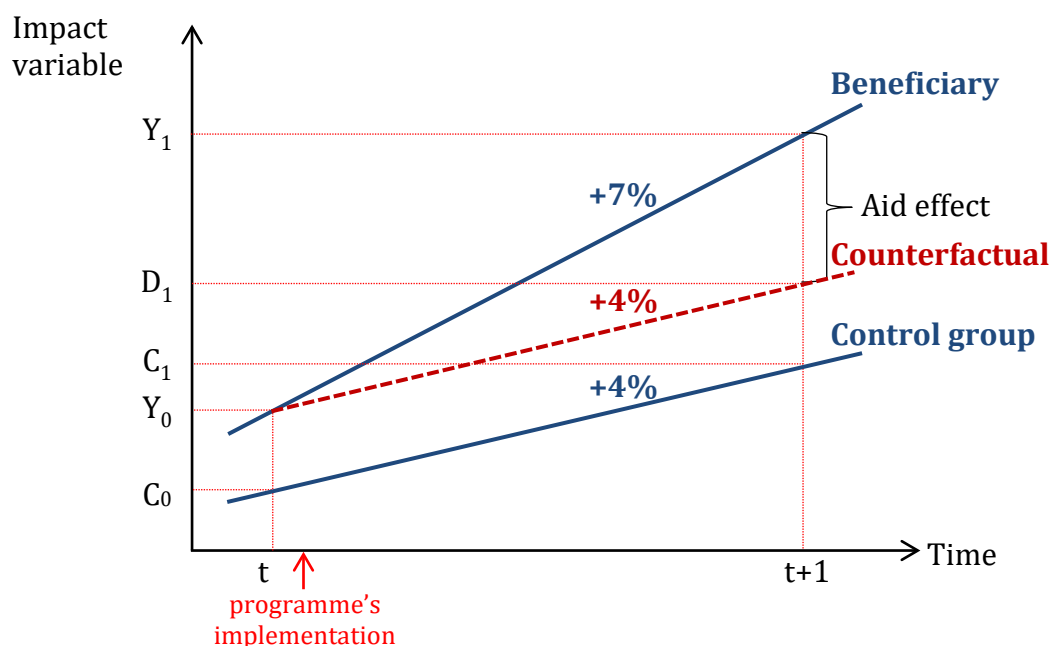
exclusively to the aid scheme instead of looking into possible *ex ante* dissimilarities.

Yet, the comparison between beneficiaries and non-beneficiaries is used in some evaluation techniques like the double difference. As will be explained below in the section dealing with that method, some caveats must be taken into account when following that approach.

These examples show the reason why **ex post impact evaluation** is so daunting. The exercise must compare an observable outcome (what actually happened to beneficiaries) with an unobservable scenario like the counterfactual (what would have happened to these same beneficiary firms in the absence of aid). For that purpose, the knottiest issue is **building a control group of firms comparable to the beneficiaries, serving as a reliable counterfactual** (Khandker *et al*, 2010). In the abovementioned examples, the chosen counterfactuals (the same firms in previous moments or other firms in the same moment) were not necessarily valid because of the reasons already discussed.

Figure 4.3 represents how impact evaluation can be carried out if we are able to find a good control group of firms, comparable to the beneficiaries. If in the control group the impact variable has risen a 4% (from C_0 to C_1), we can assume that the same would have happened to the beneficiaries in the absence of aid (the counterfactual), experiencing as well a 4% increase (from Y_0 to D_1). Consequently, the effect of the aid scheme ($Y_1 - D_1$) is the difference between the actual outcome (Y_1) and the counterfactual (D_1 , so that $D_1 = Y_0 + [C_1 - C_0]$), an extra 3pp in the growth rate (from 4% to 7%).

Figure 4.3. Choosing a valid counterfactual (control group)



The cornerstone of evaluation lies in finding a valid control group to construct an appropriate counterfactual. There are several methodologies to address these issues, although sometimes the choice of the technique is determined by data availability concerns. The **alternative approaches** can be divided into two groups:

- **Experimental design (randomization):** this technique can only be applied when programme's beneficiaries have been selected totally at random. In those cases, evaluation is relatively straightforward given that, if the sample of firms is sufficiently large and is appropriately designed, there should not be systematic *ex ante* differences between beneficiaries and non-beneficiaries. As a consequence, *ex post* differences could be attributed to the aid itself. In other words this is the only case where, in principle, every subset of non-beneficiaries is a valid control group to build a counterfactual. But experimental design is not free from pitfalls, especially the debatable rationale for allocating public funds randomly.
- **Quasi-experimental techniques:** as experiments are not frequent with public policies (even if increasingly common), when the aid is not granted randomly alternative methodologies use several tools to somehow try to simulate an experiment. All these techniques (propensity score matching, double differences, regression discontinuities and instrumental variables) share the spirit of exploiting information at hand to construct a control group of firms which can be a valid 'counterfactual' to estimate what would have happened to the beneficiaries without the aid and to be compared with beneficiaries' actual outcomes.

IV.2.a. Randomization

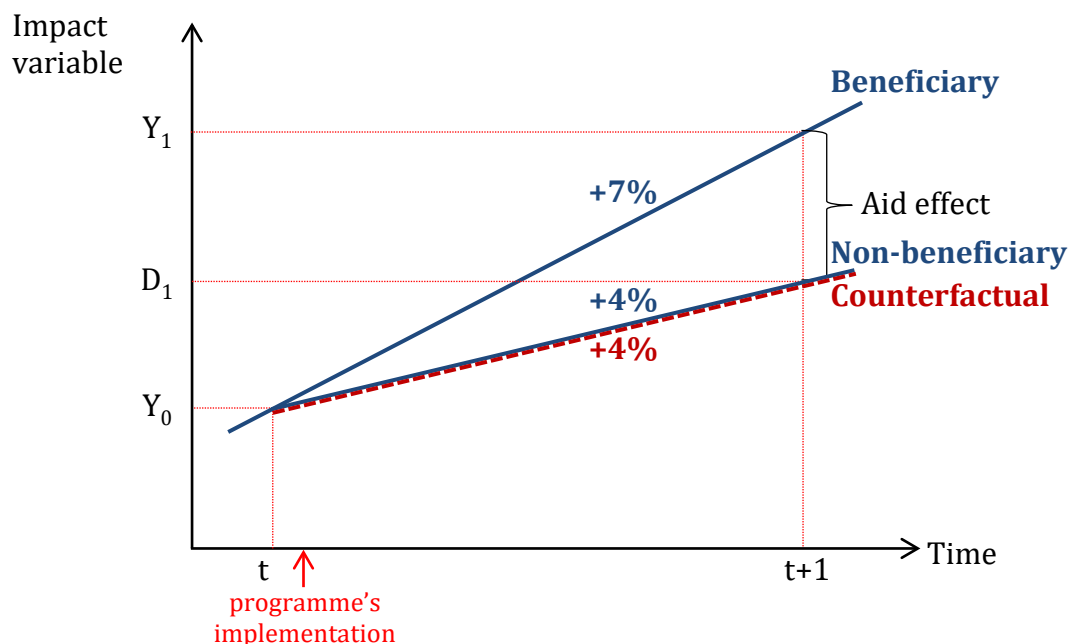
The idea of studying a given economic policy through a randomized control trial (RCT) comes from sciences like medicine and it is being recently and increasingly applied to the evaluation of development policy and aid regimes to individuals. This is not yet so common in the field of State aid, but some interesting conclusions can still be drawn from the field of development (like microcredits, which can provide lessons for aid to SMEs and entrepreneurship, as box 4.1 describes).

As far as State aid is concerned, **granting funds randomly would allow to address the selection bias**, i.e. the fact that *ex ante* dissimilarities explain programme's participation and make difficult to ascertain if (and to what extent) *ex post* divergences in outcomes are due to the scheme itself or to those *ex ante* differences. If aid beneficiaries and non-beneficiaries are selected purely at random there should not be systematic *ex ante* dissimilarities between the two groups and the whole divergence in outcomes should be explained by the aid.

Figure 4.4 represents how impact evaluation is actually implemented in this setting. If the sample is properly constructed, there should not be *ex ante* differences and beneficiaries and non-beneficiaries should depart from a similar situation (Y_0). As non-beneficiaries are a valid control group to simulate the counterfactual,

variations in outcomes (the breach between Y_1 and D_1) can be explained by the aid itself (an extra 3pp in the growth rate, from 4% to 7%).

Figure 4.4. Evaluation with an equivalent control group (selected randomly)



Yet, this approach is not exempt from **caveats** (Khandker *et al*, 2010, Moral and Pérez, 2015). The first issue is the **internal validity** of the results, i.e. whether a straightforward comparison between beneficiaries and non-beneficiaries gauges precisely the programme's impact. In order to ensure internal validity, it is therefore crucial an appropriate sample design, sufficiently large to address drivers of firms' heterogeneity. In the former figure, we assumed that enterprises departed, on average, from identical values in the outcome variable, but this might only be the case of an ideal experiment. And even starting with equal figures does not imply that firms are identical, as one should not pay scant attention to other factors, observable (like previous trends) or unobservable (such as management).

Internal validity of the experiment is thus threatened by many drivers of firms' heterogeneity like the macroeconomic environment, market potential, size, efficiency, location, financial resources, credit constraints, managerial skills or even the fact of having been granted other aid funds. These *ex ante* differences matter especially when they may cause a given firm to decide whether to carry out a project even before receiving the aid (European Commission 2014b).

Consequently, given this heterogeneity, the robustness of the conclusions may be reinforced by using econometric techniques in order to control for observables and filter them out. Moreover, heterogeneity among firms provides also a good reason to go beyond average effects and try to estimate variations in impact for different types of firms.

There are two more issues which may affect internal validity, one technical and one of economic substance (Heckman and Vytlacil, 2005). The former refers to the decrease of sample quality over time due to lack of compliance (beneficiaries which do not actually carry out the project appropriately) or selective attrition (beneficiaries and non-beneficiaries which fail to report data or just cease to exist because of factors related to the aid). The latter is the presence of spillovers amidst a general equilibrium environment where economic agents are interrelated, implying that non-beneficiaries may be indirectly affected by the aid due to changes in the beneficiaries' conduct.

The second source of concern with RCT is **external validity**, i.e. whether the results can be generalized to the whole population or even to other environment in terms of regions, sectors or time frames. There will be always a trade-off between internal and external validity: the more similar the beneficiaries and the control group, the more robust the results but the more difficult to extrapolate to different circumstances. Combining internal and external validity would require within-group randomization in all the above-mentioned drivers of firms' heterogeneity, increasing the complexity of the experiment (if not turning it unfeasible).

This shortcoming of RCT regarding external validity leads some authors to favour structural models (Deaton, 2009). As has been mentioned above (in the chapter of *ex ante* quantitative evaluations), structural models are based on somewhat universal assumptions about firm's behaviour which may be more easily generalized. Furthermore, they add value relative to RCT by giving theoretical explanations about the channels through which the aid scheme is effective or not.

But, finally, the **main issues with RCT are ethical. Granting public funds totally at random does not seem efficient *a priori***, so it is certainly difficult to justify. **However, implementing sizeable aid schemes year after year without evaluation is neither more efficient nor more ethical.** At least randomization may serve the purpose of evaluating a programme in order to see whether it should be honed, remodelled or just abandoned in order to improve public and general efficiency. Random allocation could be introduced gradually in the most market-oriented schemes, in terms of instruments (like soft loans or guarantees in contrast to grants) and of sectors (such as entrepreneurship in contrast to social schemes).

Furthermore, the random allocation of public funds may be logical and fair in some cases. For instance, in cases of oversubscription and scarcity of public funds, where requests exceed the programme's resources, allocating funds at random could be sensible in terms not only of evaluation (as non-beneficiaries would be a valid control group) but also of transparency. A randomized phased-in could also be considered according to the so-called 'pipeline methods', so that eligible population enters the program gradually, period by period (Khandker *et al*, 2010, European Commission, 2014b). Thus, during the first periods, the ones who have not benefited yet from the public policy can be a control group for evaluation purposes. This would help to assess whether the programme might need changes in subsequent periods, before being applied to the remaining eligible population.

Box 4.1.a. Evaluation through randomization

Banerjee et al (2015). An evaluation of a randomized lending microcredit programme in India showed that the availability of loans increased their take-up mildly (by 8.4 percentage points). Investment and profits increased, although these positive impacts were mostly concentrated in the *ex ante* best performing enterprises. There were as well increases in the labour supply in the own firm but there were not significant changes in business creation rates. Consumption (except for the increase in expenditure in durable goods) and welfare did not increase significantly.

Box 4.1.b. Hypothetical evaluation through randomization

Let us assume that we want to assess the effectiveness of a State aid scheme which provides subsidized training by checking whether (or not) beneficiary firms increase productivity and employment. If we select 1,000 companies totally at random and distribute funds to 500 of them (while the other 500 receive nothing) then, *a priori*, beneficiaries should be on average very similar to non-beneficiaries, so that the latter do constitute a valid control group in principle. In other words, in this case we can actually compare differences between beneficiaries and non-beneficiaries in order to estimate aid effects.

The internal validity of this experiment would only be threatened if the sample is not properly designed (for example, if there are too many large companies that distort aggregate results within the 500 companies chosen at random), something unlikely if we have many observations. Results would also lack robustness should there be problems with data collection, for example, if certain beneficiary companies fail to report data (this would normally be the case of worse-performing companies, so the actually collected data could lead to an excessive optimism, since beneficiary firms which are still reporting would happen to be the best performers). The question of spillover effects must also be heeded, given that if aid were to be effective beneficiaries would increase its productivity and employment, propelling as well non-beneficiaries' productivity and employment (misleading us into concluding that the aid has no impact if beneficiaries and non-beneficiaries improve to the same extent).

The issue of external validity is usually more relevant in experiments. In other words, the results we get from the evaluation are valid only for this type of firms under these space-time conditions. But we do not know how our training programme would work at a different juncture (like a recession) in another region (where companies tend to be larger or more export-oriented) or in other kind of companies (for example, if we aimed to focus on very small firms). In order to extrapolate our conclusions to other types of firms (improving external validity), we would have to generate more random samples to take into account all these factors of heterogeneity (increasing the complexity of the experiment and impairing internal validity).

Nonetheless, one should not resort to the issue of external validity mechanically in order to hinder an exercise of evaluation or the application of its results. For example, if we have multiple studies pointing to some flawed designs of training programmes, we should take them into account for our scheme instead of unceasingly argue that conclusions cannot be extrapolated (because the evaluation results were obtained in a different region or period).

IV.2.b. Propensity Score Matching

The allocation of most aid schemes is not made at random. On the contrary, programme participation will depend on some intrinsic differences between beneficiaries and non-beneficiaries, making extremely difficult to disentangle the programme's actual impact from those *ex ante* factors. Yet, there is still the possibility to **estimate aid effects by comparing participants to those non-participants which are very similar to them in many observable**

characteristics (except from the fact that the former did receive the aid and the latter did not).

This technique is called Propensity Score Matching (PSM) and must be framed in three steps (Khandker *et al*, 2010):

- 1st. A model is constructed to estimate the probability of being a beneficiary of the regime as a function of several observable features: turnover, experience, location, etc. In order for the technique to be robust, the model should not leave apart any other relevant aspect influencing programme's participation. In addition, it is crucial to ensure that these variables are not affected by the aid scheme.
- 2nd. Relying on the previous model, the probability to participate (*propensity score*) can be computed for the whole bunch of firms. The exercise will focus on those firms having a similar estimated propensity score. It is essential to have a sufficiently wide sample, with a balance between beneficiaries and non-beneficiaries in this region of overlap or common support. Therefore, there will be a trade-off (known as the *dimensionality curse*) between this second phase and the first one. If one chooses more variables to estimate the probability of participation, this propensity will be measured more precisely but the base of comparable beneficiaries and non-beneficiaries will be smaller.
- 3rd. Firms whose probability is relatively close can be considered similar on average, except from the fact that some did participate in the programme and the rest did not. Therefore, the scheme's impact can be estimated by matching participants to similar non-participants, attributing *ex post* differences to the aid itself. This matching can be done through several tools (Khandker *et al*, 2010, Moral and Pérez, 2015):
 - The easiest technique is the straightforward comparison of a beneficiary's value with its nearest non-beneficiary neighbour(s) or with non-beneficiaries within a given range. The matching can be done without or with replacement (i.e. removing or not a non-participant whenever it is used for matching). These effects computed at the level of individual beneficiaries must be aggregated according to their relative weights to find the programme's overall impact.
 - More general tools may include matching by intervals or *strata* (which, again, should be aggregated according to their relative weights to find the programme's overall impact) or non-parametric techniques (by using all observations to construct estimates of firms' distribution across the different values).
 - Double difference matching estimators can also be constructed (Asdrubali and Signore, 2015), computing the programme's impact as the differences in changes for beneficiaries and non-beneficiaries. As we will in the section devoted to double difference, these estimates filter out not only observable

factors but also unobservable factors explaining programme's participation (provided that they are time-invariant).

This technique is very useful, not only to check to effects of the aid itself (see box 4.2) but also to look into the factors which explain programme's participation, which can be exogenous (due to conditions set by the granting authority in terms of firm's age, size or location) or endogenous (given that some kinds of firms are more likely to apply for and receive State aid).

Box 4.2. Evaluation through propensity score matching

Whitacre et al (2014). This exercise does not study an aid regime but it does look into the effects of broadband deployment in the US between 2001 and 2010 under a PSM framework. US Counties with high broadband penetration were compared to a group of control of similar counties with lower broadband. In order to build this comparable group, the probability of broadband adoption is estimated as a function of several socioeconomic factors (income, education, firms, metropolitan areas...). Afterwards, counties with a similar probability but with actually different deployment are compared, finding that broadband adoption meant higher income growth, lower unemployment and more firm creation. Impacts in terms of broadband availability (in contrast to adoption) were more modest, implying that policies should rather be demand-oriented.

Criscuolo et al (2012). This paper evaluates the effects of regional development policy in the UK. Among other techniques, it uses propensity score matching to estimate a probability of participation as a function of several observable variables (labour productivity, employment, intermediate inputs, capital, age and a multinational dummy). The authors suggest that matching by area could also be used, by looking at ineligible areas closer in observed characteristics to eligible areas, and then comparing similar plants and firms between regions. The paper also tries other approaches, like instrumental variables. Overall, the results show that a 10% investment subsidy triggers a 7% increase in employment, coming equally from existing (normally small) firms and from newly created businesses. This boost in occupation generates a reduction in unemployment in beneficiary regions, without coming at the expense of other regions' (rejecting any risk of crowding-out).

IV.2.c. Double differences

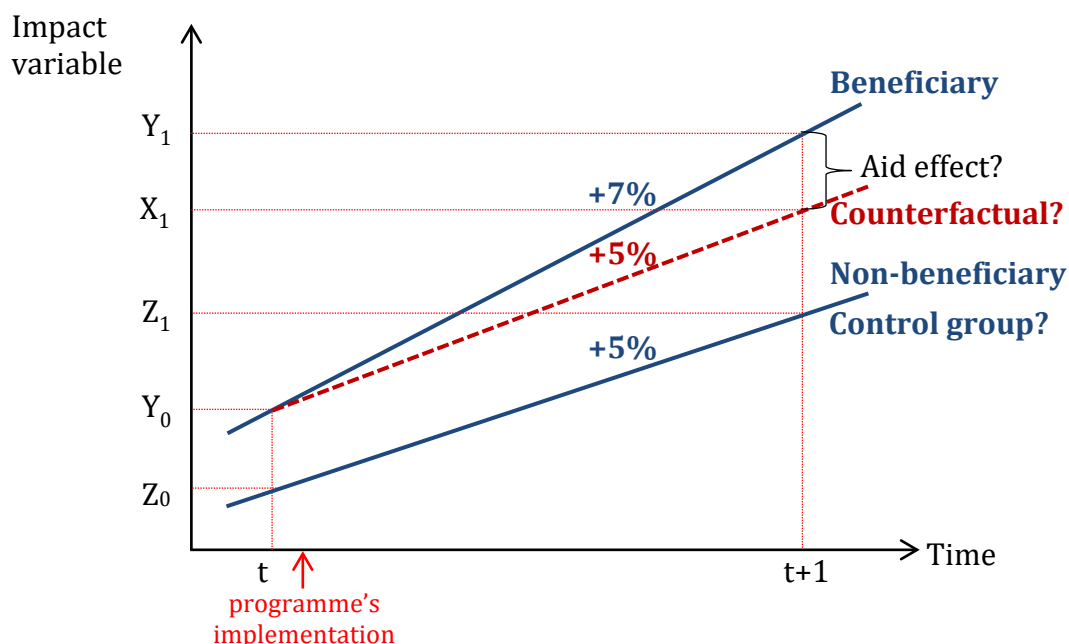
Double difference (DD) methods **compare beneficiaries with non-beneficiaries in order to capture the aid impact**. As was said above, this approach has to take into account some assumptions and caveats. Otherwise, it could lead to flawed interpretations.

Figure 4.5 (virtually identical to figure 4.2.c) shows how evaluation is implemented within this framework. It is called 'double difference' because it requires the computation of two differences (Moral and Pérez, 2015):

- The first difference is the change in the impact variable for beneficiaries ($Y_1 - Y_0$), due both to the aid scheme and to other factors, and for non-beneficiaries ($Z_1 - Z_0$), explained only by factors other than the aid regime.
- The second is the difference between those two differences (that is why the method can be named 'difference in difference' too), which is the estimate of the programme's impact (given that $[Z_1 - Z_0] = [X_1 - Y_0]$):

$$(Y_1 - Y_0) - (Z_1 - Z_0) = (Y_1 - Y_0) - (X_1 - Y_0) = (Y_1 - X_1)$$

Figure 4.5. Evaluation through double differences



The **main caveat** is the **resort to non-beneficiaries as a control group in order to build a counterfactual to which beneficiaries can be compared**. The advantage of this method is that **it does not require beneficiaries and non-beneficiaries to be homogenous**, as **ex ante** differences can be cancelled out by this double difference approach. But one does need to assume that these differences are time-invariant and unrelated to programme's participation.

These assumptions are tricky and could threaten the results' internal validity. Differences between beneficiaries and non-beneficiaries are not necessarily static and could be widening regardless of the aid effects. At the same time, when these *ex ante* differences are related to programme's participation, it will be challenging to disentangle *ex post* dissimilarities (owing to the aid itself) from *ex ante* ones (which already existed before the programme's implementation). For instance, a change in the firm's management could explain both the fact of being granted aid funds (because the new managers are more aware of these opportunities) and an improved performance. In all these cases, the aid effect calculation in figure 4.5 would overestimate the actual impact (in other cases there could be underestimation).

In order to control for these factors, the DD approach could be inserted in an econometric model that takes into account time-trend effects and pre-programme conditions and characteristics (Khandker *et al*, 2010). Furthermore, the comparison between beneficiaries and non-beneficiaries should be analysed before the aid too, in order to assess whether the latter are a valid control group or whether dynamics were already diverging before the aid implementation (European Commission, 2014b).

That is why, as was said before, the PSM method could be combined with DD (see box 4.3). PSM could work without *ex ante* data but it does require a model to estimate programme's participation as a function of all relevant observable variables. Yet, the PSM is not valid whenever there are unobservable differences between beneficiaries and non-beneficiaries.

Meanwhile, DD does allow for unobservable differences between participants and non-participants, as long as they are time-invariant. But DD does need *ex ante* data to calculate changes in outcomes and these may be difficult to collect especially for non-beneficiaries.

There are other strategies to follow when applying this technique (European Commission, 2014b). For example, DD is more precise when the lapse of time for evaluation is relatively short, because the risk that beneficiaries' and non-beneficiaries' paths diverge substantially is lower.

At the same time, DD is more accurate when non-beneficiaries did not participate because they were not eligible. In other words, we rule out the case where programme's participation decision depends on endogenous factors, i.e. beneficiaries are granted State aid because they already had a profitable project while non-beneficiaries were *ex ante* worse performers (and did not apply or did not receive the aid because of that). That would be the prototypical case when DD is not valid *a priori*, as *ex post* divergence could be due to both the aid itself and to the *ex ante* dissimilarities related to programme's participation. But if non-beneficiaries did not participate because of an exogenous factor set in the scheme, like size or location, then they can be an appropriate control group to which beneficiaries can be compared (as beneficiaries and non-beneficiaries can include in the same extent good performers with profitable projects to be carried out).

Nonetheless, even when eligibility is conditioned on firms' dimension or geographic area, comparison using DD has its pitfalls, since it will necessarily comprise companies of distinct size or location, which are intrinsically different. For instance, if we take a support scheme for SMEs in a given region, we have two possible but imperfect control groups: non-SMEs in the same region (which face similar economic conditions but are different because of size) and SMEs in a neighbouring area (which are similar in terms of size but could face a different environment). In a case like that, a triple difference approach can be followed (European Commission, 2014b) to estimate the programme's impact sequentially:

- The first difference is the change in the impact variable for beneficiaries (due both to the aid scheme and to other factors), SMEs in our region, and the change in the impact variable for non-beneficiaries (explained only by factors other than the aid regime), non-SMEs in our region.
- The second is the difference between the former two differences (the 'difference in difference'). This could be the first tentative estimate of the scheme's impact, but it is biased due to the fact that we SMEs have idiosyncratic features with respect to non-SMEs (the control group). That is why we make use of a third difference.

- We compare the change in the impact variable for SMEs and non-SMEs for the neighbouring region(s). These are unaffected by the programme (none is a beneficiary) but shed light on the difference between SMEs and non-SMEs. By comparing the second difference (the former biased estimate which compares SMEs and non-SMEs in our region) with this other difference (the comparison between SMEs and non-SMEs in neighbouring areas) we get the triple difference (the ‘difference in *difference in difference*’) as a more precise estimate of the programme’s impact.

Box 4.3. Evaluation through double differences

Asdrubali and Signore (2015). This paper estimated the impact of an EU SME guarantee (SMEG) facility in Central, Eastern and South-Eastern European (CESEE) Countries in the period 2005-2012, combining PSM and DD. Firstly, the probability of participating in the scheme was estimated as function of several factors, finding a positive relationship with profits and fixed assets and a negative relationship with liquidity and the number of employees. Secondly, beneficiaries were matched to their closest non-beneficiaries (with a very similar estimated probability to participate), following a double difference approach, i.e. the change in beneficiaries’ values were compared to the change for non-beneficiaries’. Although results were country-specific, the impact was mostly positive, with an increase of 17.3% in employment and of 19.6% in turnover for beneficiaries (relative to the control group of similar non-beneficiaries). These effects were even more sizeable for smaller and younger firms. However, profits did not grow and productivity and liquidity actually worsened in the short term (showing at least some concerns to digest new investment projects).

IV.2.d. Instrumental variables and regression discontinuity

As we have said, the **trickiest issue when carrying out an evaluation is endogeneity. Some variables which influence participation in the aid regime also affect outcomes**, so it will be impossible to determine which part of the change in those outcomes could be attributed to the aid itself. When endogeneity is not taken into account, impact estimates are biased and hence unreliable.

Instrumental variables (IV) try to overcome this setback by **seeking a variable** (the ‘instrument’) which is **correlated with a programme’s take-up but not with its outcomes** (Khandker *et al*, 2010, European Commission, 2014b). This is done through a two-stage approach:

- 1st. Programme’s participation is estimated as a function of this instrument and other relevant variables. If the instrument is not statistically significant, one should look for another valid instrument or try alternative specifications with other relevant variables.
- 2nd. The effect on the impact variable is estimated as a function of the fitted value of programme’s participation obtained in the first stage (it should be closed to actual participation if the instrument is good enough) and other relevant variables.

The IV approach has several **concerns**. Chief among them is **the challenge of finding a good instrument**, which has to be strongly correlated with participation in the aid scheme but not at all with the outcome variable.

Regression discontinuity (RD) is a particular case of IV which takes advantage of **eligibility criteria as exogenous factors** (like firms' size or experience) which **determine access to aid funds** (Moral and Pérez, 2015). As these variables are correlated with participation in the scheme, they can be *a priori* valid 'instruments'. Ideally, these exogenous factors should be continuous at both sides of the threshold which introduces the discontinuity.

There are two types of RD design:

- **Sharp design:** the correlation between the instrument and participation is perfect (i.e. deterministic), because firms above or below a given threshold are beneficiaries and the rest are not.
- **Fuzzy design:** the correlation between the instrument and participation is high but not perfect (i.e. stochastic), so most beneficiary firms are above or below a given threshold but there could be as well beneficiaries below or above that level.

RD is a specific case of IV. Programme's participation has to be estimated as a function of this exogenous factor and other variables and then this estimate of programme's participation is used to assess the effects on outcomes. In sharp design, where the relation between the instrument and participation is perfect, the first phase can be skipped and one can compare directly outcomes for agents just below and just above the threshold (Khandker *et al*, 2010). This looks like the PSM too, to the extent that it aims to compare firms which are very similar in everything except for the fact of being granted funds, so that the impact of the aid itself can be captured.

And RD, apart from being an application of IV with points of resemblance to PSM, also tries to replicate RCT. With the additional advantage that, while randomization is quite uncommon in the scope of State aid, RD can be easily applied because there are many regimes which use thresholds (like a given firm's level of turnover, employment or years of experience) as these types of criteria to allocate funds are endorsed by public opinion, in contrast to randomization.

Indeed, RCT and RD are sometimes combined in the abovementioned 'pipeline methods' (Khandker *et al*, 2010), which take advantage of timing and delays in programmes' implementation to compare firms which have already been given the funds with those which are equally eligible but have not yet received the aid (they are 'in the pipeline'). This sequential approach is endorsed by the public too (given that public funds are scarce and sometimes they cannot be distributed at once to the whole eligible population) and is a good groundwork for evaluation if this gradual access is randomized.

But **RD has its shortcomings too**. In principle it should be subject to the same restriction as the IV in the sense that exogenous factors determining eligibility (the 'instrument') should not be correlated with the outcome variable. That would be challenging given that the most frequent eligibility criteria seem to be correlated with outcomes. For instance, when aid schemes pick small or young firms, these

are bound to be intrinsically different from non-beneficiaries (bigger and more experienced firms), so it will be impossible to disentangle the actual impact of the aid from these other *ex ante* dissimilarities by merely comparing beneficiaries and non-beneficiaries.

RD normally addresses this issue by **analysing only firms at both sides which are very close to the threshold**. This allows relaxing the assumption of no correlation between the instrument (the eligibility criteria) and the outcomes, because the effects of that correlation would be small given that companies have still very similar values of that instrument (European Commission, 2014b).

But this strategy has two additional setbacks (Khandker *et al*, 2010). On the one hand, there could be only a few firms close to the threshold so the **sample of comparable enterprises would be reduced** to a point where it would be difficult to get a statistically robust result (i.e. internal validity would be jeopardized). On the other, even if finding convincing effects, these would be applicable only to the companies around that threshold and **might not be extrapolated to other types of firms** (i.e. external validity would be impaired as well). That is why RD is also called 'randomization in the bubble' (Moral and Pérez, 2015), because it aims to mimic a randomization setting but it does so in a very specific environment, the 'bubble' of firms which are very close to the threshold (even if at both sides of it).

Box 4.4. Evaluation through instrumental variables and regression discontinuity

Becker *et al* (2010). This exercise takes advantage of EU regional policy, which provides funds to NUTS-2 regions with a *per capita* GDP level below 75% of the EU average, to develop a RD. The design is fuzzy, because regions over that threshold receive some support, albeit smaller. Therefore, the 2 stage conventional IV approach is followed. First, the probability of receiving funds is estimated as a function of pre-programme income. Secondly, the impact of that probability on outcomes is checked. The authors find positive *per capita* GDP growth effects but no significant changes in employment.

V. Concluding remarks

State aid can be a powerful tool to actually improve efficiency and general welfare. But it has to fulfil **two conditions**:

- i. There is a **market failure** (a situation where the market does not lead to an optimal outcome owing to externalities, public goods, imperfect information or increasing returns) or **another objective of common interest** (such as redistribution) to be addressed. This is a necessary condition for government intervention (through State aid or other means) but not a sufficient one. That is why we add the following prerequisite.
- ii. The principles of **necessity, appropriateness, proportionality and minimum distortion** must be respected. In other words, there is a market failure or an objective of general interest (*necessity*) and State aid is the most suitable measure to tackle it (*appropriateness*) in comparison with other government policies (such as taxation or regulation) or with solutions within the market, including the possibility of no intervention. Furthermore, the aid scheme's benefits (given the magnitude of the market failure or objective targeted) must exceed its potential costs (*proportionality*) and the design of the scheme must keep these costs at the lowest possible level (*minimum distortion*), considering alternative designs of the regime.

This caution in granting **State aid** is warranted because of its potential **shortcomings for competition, efficiency and public finances**:

- ✗ The **free play of competitive forces might be altered** by an intervention that gives resources to a selected group of firms. Furthermore, State aid can create or intensify contexts of entry barriers, market power, consolidation and collusion.
- ✗ **Efficiency could be jeopardized** by influencing the allocation of resources by the market. Resources could be diverted towards inefficient conducts (like lobbying for State aid) and firms lose incentives to seek the most efficient technology. Finally, trade and investment flows could be distorted.
- ✗ **State aid has an obvious cost for public finances**, as resources come from higher taxation and/or lower expenditure in other areas which could be more productive. This issue is well beyond the scope of a competition authority, but it is relevant to increase the dose of prudence when assessing aid schemes. Even when an aid regime has a positive balance from the point of view of efficiency and competition, it would not be necessarily applied if it is less welfare-enhancing than other measures which compete for the scarce public resources.

Therefore, **State aid schemes have to be evaluated** to ensure that the former conditions and costs have been taken into account.

An exercise of **ex ante evaluation** should take place before the programme is being adopted:

- **Every State aid regime should be the object of an *ex ante* qualitative evaluation** to safeguard the criteria of **necessity, appropriateness, proportionality and minimum distortion** (the ‘State aid balancing test’). For instance, when a programme lacks a convincing justification of the market failure or the objective of general interest to be addressed, this should be a warning signal to reconsider it (because there are actual risks that the scheme is the result of ‘subsidy races’, that want to attract economic activity in the short-term, or is influenced by political pressures and lobbies). At the same time, a cost-benefit analysis and the comparison with other alternatives should be carried out before adopting the aid programme.
- **Some State aid regimes could be evaluated as well using *ex ante* quantitative evaluation methods**, such as econometric estimations, whose results could be extrapolated to our case, or structural macroeconomic models, whose theoretical framework can fit our situation (if abstractly). This approach can be followed for **wholes lines of policy** or for schemes of a **relevant size** within the granting authority’s budget.

An exercise of ***ex post* evaluation** ought to take place after the programme has started being implemented (and is still ongoing or has already ended):

- **Every State aid regime should be subject to monitoring and operational evaluation** to check whether final objectives and intermediate targets were met or not. This is a compulsory first step in transparency, accountability and policy learning.
- **Some State aid regimes could be assessed as well using *ex post* quantitative impact evaluation methods** to assess whether (and to what extent) the programme did have a direct ‘**incentive effect**’ (by triggering a change in beneficiaries’ conduct that was necessary to address the market failure or the objective of general interest) or other **indirect effects** (positive or negative, foreseen or unforeseen). In order to find a robust answer, the exercise must **compare the actual outcome following the aid implementation with what would have happened in a ‘counterfactual’ scenario without the aid**. This requires building a ‘control group’ of firms to which beneficiaries can be compared, so that the aid effects can be disentangled from other differences between beneficiaries and non-beneficiaries which should not be attributed to the aid itself. **This task involves cost and complexity**, so it should only be **applied to a limited amount of schemes**. The recently refined **EU framework** currently prescribes evaluation for **regimes above €150 million annually** (in certain fields) or could suggest it for individual **schemes which are novel**, which **affect markets whose regulation or technology is changing** or which have a **high degree of sectoral and geographic selectivity**. Needless to say, **public bodies could and should go beyond the minimum EU requirements** and **evaluation could be broadened** to other **schemes of a relevant size** within the granting authority’s budget (even if below the €150 threshold). *Ex post* impact assessment is also useful in **renewable, recurrent and permanent regimes**. And finally, evaluation might be carried out also in schemes where it is relatively straightforward given the **availability of data and information**.

The **final step** in this process involves **embedding evaluation** in the whole cycle of **policy formulation** so that **policy-makers acknowledge the relevance** of the exercise and **apply the results to actually modify** (or even abandon) **State aid regimes if needed**. Indeed, **this framework of *ex ante* and *ex post* evaluation should be applied not only to State aid but also to other public policies.**

Appendix 1. Market failures and State intervention

Normally, the free play of demand and supply forces in a **competitive market** leads to an outcome where **efficiency is maximized**. The concept of economic efficiency has three dimensions:

- ✓ **Allocative efficiency:** resources flow wherever they are most valuable, maximizing welfare. Prices of goods and services become the proper signal of relative scarcity, consumer's preferences and producer's technology.
- ✓ **Productive efficiency:** producers use the least costly and the most efficient technology to deliver goods and services, ensuring there is no waste.
- ✓ **Dynamic efficiency:** the former two static properties guarantee that economic agents are incentivised to go on seeking efficiency, as those which do not use resources efficiently are moved out of the market.

However, the former result (known as the First Fundamental Theorem of the Welfare Economics) is sensitive to several assumptions made about the context in which exchange takes place. **When** these assumptions do not hold and **the market does not lead per se to an optimal allocation** of resources, then we are in front of **market failures**, which in turn can be divided into four different kinds: externalities, public goods, imperfect information and increasing returns.

Externalities happen when the consumption/production of a good/service involves some social costs/benefits, which go beyond the private costs/benefits for the consumer/producer. In those cases, market prices are not an accurate signal of efficiency, as they reflect private but not social costs/benefits. In the case of negative externalities, like environmental pollution, the market leads to an output above the optimal (market price is below the optimal). Meanwhile, in positive externalities, like training or research, market output is below optimal levels (market price is above the optimal).

Therefore, one possibility to fix this misallocation is granting State aid for certain goods, services or technologies. There are other alternatives, such as regulation, taxation and the creation of property rights markets.

Public goods are those whose consumption is non-excludable, i.e. it is difficult to actually impede the access to them, and non-rival, i.e. the use made by an agent does not affect the use by others. This is in contrast to private goods, which are both excludable and rival. While for private goods the market generates the optimal output, public goods tend to be infra-provided. So the solution is normally for the public sector to produce directly the good or to incentivize its provision by the private sector through subsidies or regulation.

Beyond the more traditional cases of public goods (like national security or defence), a modern example could be financial stability. It is non-rival, as a bank benefiting from positive market conditions does not preclude others from doing so, and non-excludable, because a sound system allows every entity to take advantage of it. That is why there are frequently regulations which force or incentivize banking entities to follow a prudent behaviour so that they contribute to

financial stability. When this is not enough (i.e. an institution has not internalized the externality it creates by assuming an excessive risk), the state can consider granting aid to an individual agent or to the system as a whole.

But public goods in their purest form are rare and they tend to manifest in a somewhat mixed form, where the menu of options to address the market failure is wider. On the one hand, there are club goods, such as technology and not congested network infrastructures (in energy, telecommunications, transport and postal services), which are excludable but non-rival. On the other, there are common goods, such as natural resources, which are rival but non-excludable, and hence being very similar to externalities (following with the example of the environment, saying that it is a public good infra-provided by the market is equivalent to state that environment conservation is a positive externality where the market leads to an output below the optimal or, alternatively, that environmental deterioration is a negative externality where the market leads to an output above the optimal).

Imperfect information is perhaps the most frequent market failure, as agents take usually their decisions without perfect knowledge. Asymmetric information is especially harmful for market efficiency, as the side with the information advantage, known as the agent, will try to exploit it in a context of conflicting interests with the other part, known as the principal.

Firstly, there is the case of moral hazard (hidden action), where one side of a transaction has an information advantage subsequent (*ex post*) to the signing of a contract. For instance, this agent's effort could be hard or costly to monitor, if not impossible. Efficiency problems will be exacerbated by risk aversion amidst an environment of uncertainty, where the agent does not control the results obtained with its actions. So, for example, when the principal observes bad results, it does not know if they are due to a low effort by the agent or to pure bad luck. As a consequence, some transactions may not be completed, generating costs in terms of efficiency.

Secondly, there is the case of adverse selection (hidden characteristics), where one side of a transaction has an information advantage previously (*ex ante*) to the signing of a contract. For instance, the productivity of this agent may be difficult to assess). So the principal would react offering the agent a remuneration corresponding to the average productivity in the market, which, paradoxically, will be beneficial for low-productivity agents and harmful for high-productivity ones, driving quality out of the market.

In those circumstances, there is room for state intervention. The government can provide aid so that more transactions can be completed. It can also pass regulation to increase the flow of information, although the market itself can provide mechanisms to allow that exchange of information.

Credit markets are perhaps the most paradigmatic example of imperfect and asymmetric information. Previously to the signing of a contract, lenders might not be able to discriminate between borrowers on the grounds of their credit quality

(*ex ante* hidden characteristics or adverse selection). After signing the contract, lenders cannot know whether borrowers' ability to repay is affected by poor management or by the natural risk of entrepreneurial projects (*ex post* hidden action or moral hazard). Therefore, some transactions are not completed and there is credit rationing, affecting especially small, nascent and innovative firms. The government tends to intervene in credit markets with subsidies, soft loans or guarantees so that more transactions can be completed (it intervenes as well with information and registration requirements).

Increasing returns refer to the context where it is more efficient to concentrate production in a few firms because there are high fixed costs and average costs are decreasing in the scale of production. The most extreme case would be natural monopoly, where the optimal solution is for one single firm to absorb the whole demand. Increasing returns are frequent in network industries such as transport, telecommunications, energy and postal services.

Problems arise because, in these circumstances of market power, firms do not have the incentive to behave competitively. To a large extent, the incumbents can behave independently of actual or potential competitors, as initiating new production plans would be very expensive given the high fixed costs. Hence, efficiency is jeopardized in the form of higher prices and lower quantity and quality. Furthermore, if firms behaved efficiently and priced at marginal cost, they would recover variable costs but not fixed costs, incurring losses.

As a consequence, there is room for government intervention through many ways. First, the public sector should consider whether market power is the result of unjustified regulatory obstacles that hinder competition, removing those barriers should that be the case. Competition policy may also play a role where this dominant position is being abused. If market power cannot be diminished through deregulation, liberalisation and competition policy, the government could ponder regulation or State aid to ensure an adequate supply at a reasonable price.

Selected references

Asdrubali, P. and Signore, S. (2015): “The Economic Impact of EU Guarantees on Credit to SMEs. Evidence from CESEE Countries”, *European Economy Discussion Paper* 002.

Banerjee, A., Duflo, E., Glennerster, R. and Kinnan, C. (2015): “The miracle of microfinance? Evidence from a randomized evaluation”, *American Economic Journal: Applied Economics*, 7 (1), 22-53.

Becker, S.O., Egger, P.H. and von Ehrlich, M. (2010): “Going NUTS: The effect of EU Structural Funds on regional performance”, *Journal of Public Economics* 94, 578–590

Claici, A (2012): “Economic Analysis in EU State Aid Control”, Barcelona GSE Regulation and Competition Seminar Series.

CNC (2008): “Informe Anual de Ayudas Públicas”

CNMC (2015): “Informe de la CNMC sobre las Ayudas Públicas en España”

Criscuolo, C., Martin, R., Overman, H. and Reenen, J.V. (2012): “The causal effects of an industrial policy”, *NBER Working Paper* 17842.

Czernich, N., Falck, O., Kretschmer, T. and Woessmann, L. (2009): “Broadband Infrastructure and Economic Growth”, *CESifo Working Paper* 2861.

Deaton, A. (2009): “Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development”, *NBER Working Paper* 14690.

European Commission (2014a), “Draft Commission Notice on the notion of State aid pursuant to Article 107(1) TFEU”, January.

European Commission (2014b), “Common methodology for State aid evaluation”, *Commission Staff Working Document* (2014) 179.

European Commission (2014c), “You can’t improve what you can’t measure: State aid evaluation”, *Competition policy Brief*, Issue 7, June.

European Commission (2014d), “Provisional supplementary information sheet for the notification of an evaluation plan”, July.

Guellec, D. and van Pottelsberghe de la Potterie, B. (2001): “R & D and productivity growth: panel data analysis of 16 OECD countries”, *OECD Economic Studies*, 33, 2001/II.

Heckman, J. J., and Vytlačil, E.J. (2005). “Structural Equations, Treatment Effects, and Econometric Policy Evaluation.” *Econometrica*, 73(3), 669–738.

Khandker, S.R., Koolwal, G.B. and Hussain, A.S. (2010): “Handbook on impact evaluation: quantitative methods and practices”, The World Bank.

Koetter, M. and Noth, F. (2015): “Bank bailouts and competition. Did TARP distort competition among sound banks?”, *ECB Working Paper Series* N° 1804.

Kollmann, R., Roeger, W. and in’t Veld, J. (2012): “Fiscal Policy in a Financial Crisis: Standard Policy versus Bank Rescue Measures”, *American Economic Review: Papers & Proceedings*, 102(3), 77–81.

Mohl, P. and Hagen, T. (2010): “Do EU structural funds promote regional growth? New evidence from various panel data approaches”, *Regional Science and Urban Economics* 40(5), 353-365.

Moral, I., and Pérez, C. (2015): “Técnicas de Evaluación de Impacto”, Editorial Grupo Garceta.

Nicolaides, P. (2015): “*Ex post* evaluation of State aid measures”, State aid uncovered, StateAidHub.eu by Lexxion.

OECD (2005): “Denmark - Report on competition law and institutions (2004)”.

Roeger, W., Varga, J. and in ’t Veld, J. (2009): “Modelling the Lisbon Strategy: Analysing policies to promote knowledge investment with an endogenous growth model”, *Comparative Economic Studies*, 51, 520-539.

Qiang, C. Z., Rossotto, C.M. and Kimura, K. (2009): “Economic Impacts of Broadband” in “Information and Communications for Development: Extending Reach and Increasing Impact”, The World Bank.

Varga, J. and in’t Veld, J. (2011): “A model-based analysis of the impact of Cohesion Policy expenditure 2000–06: Simulations with the QUEST III endogenous R & D model”, *Economic Modelling*, 28, 647–663.

Whitacre, B., Roberto, R. and Strover, S. (2014): “Broadband's contribution to economic growth in rural areas: Moving towards a causal relationship”, *Telecommunications Policy* 38, 1011–1023.

WTO (1994): “Agreement on Subsidies and Countervailing Measures”.

