Almatourism

Journal of Tourism, Culture and Territorial Development

How Much Will this Event Benefit Our Economy? A Checklist for Economic Impact Assessments with Application to Milan 2015 Expo

Massiani, J.*
University of Milano-Bicocca (Italy)

ABSTRACT

Policy makers, policy advisers and the general public frequently perceive events as beneficial to the economy of the host city. Such belief is supported by numerous Economic Impact studies, often based on Input-Output models, which usually exhibit large positive impacts. However, research has identified potential issues in Economic Impact studies that may result in misguiding policy recommendations. Yet, no systematic presentation of such potential flaws is available to the researcher and the practitioners. To fill in this gap, this article proposes a set of criteria based on 7 categories and 32 criteria and applies them to the latest mega events held in Italy. This application supports the value of our proposed criteria and suggests that the claims based on existing studies are debatle.

Keywords: Economic Impact; Economic Impact Assessment; Mega Events; World Fair; Milano 2015

JEL: R58 (Regional Development Planning and Policy); R11 (Regional Economic Activity: Growth, Development, Environmental Issues, and Changes);

* E-mail address: j.massiani@libero.it

Cities around the world regularly compete to host major national or international events like the Olympics, International exhibitions, European Capital of Culture and so on. This participation is motivated by the prospective benefits in terms of image, community spirit, and not last, economic impacts. While it finally withdrew from the competition, the Boston Olimpic Bid, was initially supported, as were other 2024 Olympics bid participants, by positive economic prospects *"Boston Olympics will bring millions, study says" (the Boston Globe, 25 march 2015)*. In most cases, many expectations of economic benefits rely on so-called Economic Impact Studies, which are predominantly based on the Input-Output paradigm and normally produce 'large', in any meaning, numbers in favour of the events.

Yet, many scientists have shed light on the risks of improper calculation in such studies. Anomalies range from the omission of substitution effects to the use of improper IO matrices, when technical coefficients correspond to areas (e.g. nation) larger than the one used for application (e.g. region). Other anomalies may relate to excessive confidence on initial costs and incomes estimates or to improper claims on the prescriptive value of economic impact studies results. Despite repeated warnings from the economists, such practices continue to parse economic impact studies.

It thus appears useful to provide practitioners and policy advisors with a conceptual framework to evaluate the adequacy of economic impact studies. The core of our proposal resides in a list of criteria that economic impact studies should respect. Corresponding checks can be implemented internally, in the production process of economic analysis, or externally as a tool to scrutinize third part studies. To our best knowledge, no such methodological tool is yet available. We posit that such a tool, provided it is conceived in a way that supports, rather than substitutes, reflection can enhance the quality of Economic Impact Assessment and the policy recommendations that result thereof.

To discuss this methodological proposal the present paper contains the following section. Section 2 discusses why claims about the economic impacts of mega events may sometimes be flawed. Section 3 presents a set of 7 criteria, these are further decomposed for major precision in sub-criteria and 32 elementary criteria. Section 4 provides an example of implementation of the checklist to studies made for the last mega-events hosted in Italy: Milano 2015 International Exhibition and, for comparison purpose, Torino 2006 Olympics.

1. Economic Impact Assessment would benefit from reasonableness checking criteria

In this section, we present the critical aspects of economic impact studies, shown by economic analysis. We concentrate on Input-Output models and to their extension to SAM models, while the questions are a bit different for CGE models (Massiani, 2018) that represent a growing part of impact studies (Wyk et al., 2015). We also suggest that

properly conceived checklists can improve the scientific validity of Economic Impact Assessments.

1.1 Mega event evaluation are at risk of methodological flaws

In this section, we recall how economists have shown that economic impact assessment is prone to errors (Abelson, 2011; Crompton, 1995, 2006; Davies et al., 2013; Diedering & Kwiatkowski, 2016; Dwyer et al., 2000; Jeanrenaud, 2007; Késenne, 2005; Matheson, 2008b; Mondello & Rishe, 2004; Mules & Faulkner, 1996; Porter, 1999; Siegfried & Zimbalist, 2000; Tyrrell & Johnston, 2001; Zimbalist, 2010, 2015). For instances, Crompton (1995), classifies 11 misuses of multipliers. Matheson (Baade & Matheson, 2016; Matheson, 2009) repeatedly unveil distortions of economic impacts studies that can strongly overestimate the impacts. Interestingly, Hudson regresses the results of economic impact assessments on methodological features of the corresponding studies and finds that certain misuses of economic impact concepts are highly influential of the estimated impacts (Hudson, 2001).

A typical discussion on IO rests on the fact that, as used typically, they will generate, by a methodological artefact, results that are large and always positive, by construction.
On the first point, Economic impact assessments are viewed as process for generating 'large numbers' used to gain support in favour of events' organisers and interest groups linked to them. The question of the real intentions behind economic impact assessments is however beyond the scope of this paper and we will focus on more objective aspects. A milder version of these critics correspond to the 'anything goes' nature of input-output impact calculation: any expenditure, at least with methods usually implemented, will have a positive impact on the economy.
On the second point, IO results will be positive by construction, at least as they are used typically. The vector representing the event related demand shock will consist only of non-negative values, so that the final impact will always be positive. In some other applications, the demand shock vector could have some negative values. This could be the case when the analyst has considered "substitution effects". But, for events at least, these negative values will typically be smaller than other positive components of this vector. It is not possible to prove mathematically that the result will always be positive, yet we are not aware of any Economic Impact Assessment of sport and cultural events that produced a negative result (different to the works in early IO literature on other topics (Leontief, 1966) or to negative shocks). So that economic impact will always estimate a positive impact for the variables of interest like GDP or employment. Given that such variables are predominantly perceived as desirable, this will be equivalent to a recommendation in favour of the event. This contrasts with other methods that are not bound to a positive result and the corresponding recommendation.

Introducing some distinction between economic activity and welfare, other researchers claim that economic impact studies have no prescriptive value (Matheson, 2008a) : an increase in production in a given area cannot be associated unambiguously with an

increase in welfare, when externalities and opportunity costs of resources are considered. It is unsure that these considerations can invert the positive conclusions based on increased added value which routinely reach billions of euro or dollars. But a correct interpretation of impacts, as distinct from benefits, should be made by the analysts.

A more fundamental point is that impacts figures may be highly inflated. Matheson claims that some assumptions may distort economic impact quantification 'by up to a factor of 10' (Matheson, 2008a). Most of the related concerns deal with substitution effects, labelled in various ways by economists: deadweight, crowding out. Typically, researchers express concern about the consideration of local visitors' expenditures in the impact, while such expenditures should not or, at least, not entirely be considered as a source of extra activity for the area of interest. Other researches (Oxford Economics, 2012) stress that public investment for the event, such as infrastructures, can substitute to other uses of budgetary expenditures. They also mention the difficulty of such tasks:

The results do not consider the opportunity cost of public funds, which could be used to finance other projects or lower the tax burden. Whilst this is a common criticism of economic impact analysis, speculating on what the funds could have been used for involves conjecture.

Whatever these difficulties, thus, the estimated economic impact should be net of the loss of activity resulting from reduced expenditures in other parts of the economy. The possible objection that 'this money would not have been spent without the event' is usually unsupported and does not seem consistent with private and public budgetary practices in most developed economies. Ultimately, if the claim is that the money would not have been spent without the event, this overarching assumption should be explicit and carefully motivated.

The examples we report do not imply that all potential mistakes systematically generate overestimates of the benefits. True, the institutional context, in which these studies see the light, is such that overestimates would be, even in good faith, gratefully welcome by stakeholders. However, a check-list should avoid a priori, whether positive or negative and should be as neutral as possible and consider equally issues whether they generate over or underestimation.

In conclusion on this point, some concerns appear on the risks of improper implementation of Economic Impact Assessment, mostly IO, that could inflate or reduce the estimated impacts. It then appears necessary to provide some guidance for practitioners and for potential users of such analysis who want to assess their validity. In the next section, we present a conceptual framework for checking the reasonableness of Economic Impact Assessment. Additionally, this conceptual framework can be used as a check-list for practitioners. Frame 1 presents in detail how such a checklist, as used in other professional contexts like, for instance, engineering can help reduce risks of flawed analysis or improper action. Readers that are already convinced of this point may want to proceed to the next section.

Frame 1 – Reasonableness checking or checklist, useful tools to support complex tasks

We first consider a category of documents known as Guidelines. Some **guidelines** have been issued for event economic impact assessments (Nicolas, 2007). These provide useful indications for professionals, but their realm could be limited for two reasons. First, their authority often derives from their endorsement by an administration or a professional organisation. This can be a strength but also a weakness as some analysts could just argue that the guidelines provided, say, in Scotland, may not apply south of the Adrian wall. This also means that the prescriptive power of these guidelines are contingent on the support of the authority that issued them, a support that is not guaranteed in the medium to long term. The second limitation is that they do not have the immediateness of checklists which can provide, with minimal time requirement, an evaluation. It then appears legitimate to propose, additional to existing guidelines, other tools to assist in the delivery of proper economic impact assessments. These can be referred to as checklists or reasonableness checking.

Checklists and **reasonableness checking**, as used in various fields like transport planning (TMIP, 2010) are not fundamentally different. One possible difference is that the latter usually provides benchmark quantifications that can be used to check the input or output of a given analysis. Such quantitative references could be useful for economic impact assessment, but they are beyond the scope of this article, and the question on whether economic impact assessment is a sufficiently consolidated area to allow for such approach is still an open one. This suggest that Economic Impact Assessment are mostly mature for checklist.

In parallel to this, we observe that checklists are proficiently used in a number of fields, many linked with human factor psychology or with ergonomics (Drury, 2006). In a scientific essay, properly dedicated to checklists, Gawande states that no matter how expert one may be, well-designed checklists can improve outcomes. The author provides evidence of efficiency for the simplest instructions such as the '*wash your hand*' in surgery (Gawande, 2011). McLaughlin argues: '*The use of checklists is a primitive yet remarkably effective strategy for ensuring accuracy in complex tasks*' (McLaughlin, 2010). One may however wonder whether such successful outcomes may transfer outside of the surgery activity (Haynes et al., 2009) not to mention the iconic aeronautic example (Degani & Wiener, 1990). Actually, such transfer may not be adequate: surgery checking takes place in operational context where time available for correcting actions is very limited; they typically imply various operators with very heterogeneous professional competencies. Both these conditions do not hold in typical economic impact studies professional contexts.
Yet, other fields of human activity, which are more comparable to ours, have successfully implemented checklists procedures. For instance, Chang et al have quantified the benefits of checklist for building engineers (Chang et al., 2012). It thus appears that checklist can be fruitful for assisting the proper implementation of economic impact analysis.

Checklists also come with different flavours. For instance, some checklists are based on '*if,… then do*' logic. Others, which may more closely adhere to the checklist literal meaning, merely check whether some conditions are verified, irrespective of possible counteractions. We argue that 'if .. then do' checklists are more frequent and adequate in operational fields where quick action is necessary (e.g. surgery). This may not be appropriate for economic impact assessment that usually allows for more time for reflection.

One should ask himself **what is a useful checklist.** Actually, guidelines for checklist are available, although very few, to our best knowledge, in peer reviewed contents. K. Radeka published an unreviewed 'A Checklist for Designing a Checklist".

Eventually, a limitation one may see in checklists is the risk of an excessively mechanistic use. These tools should be a support for reflection and not a substitute for it. By chance, the kind of errors that take place in economic impact assessment usually relate to complex issues, conceptual misconceptions, where mechanical, unreflected, corrections are unlikely to be feasible. Moreover, there can be more 'mechanism' in the production process that governs 'unchecked' economic analysis than in the application of checklists. For instance, a typical sequence: 'define a stimulus vector, compile an IO matrix, compute the impact using a matrix algebra, present the results', leaves little room for methodological reflection apart from some statistical and national accounting issues in the construction of the matrix.

It thus appears that properly defined checklists, designed to invite for reflection and not to substitute for it, avoiding a too mechanistic approach, are beneficial for the evaluation of events impacts.

As a conclusion for this section, it appears that many economic impact studies are at risk of providing flawed outcome. Checklists, as long as applied as a support to, rather than a substitute for, reflection are a useful tool to make these outcomes more realistic.

1.2 *Four ex ante studies about mega events in Italy*

It can be interesting to illustrate the need of checklists considering a sample of impact studies. To do this, we selected recent large events in Italy. We focus on the economic impact studies that materialises in 3 studies for Milan 2015. We also include one study for Turin 2006 Olympics to provide comparison with the previous mega event held in Italy. The material analysed for each study is presented in Table 1.

Table 1: Sources of information used for Italian mega event impact studies[1]

Event	Study	Available documents
Milano 2015	Dossier de candidature 2007.	• Comitato di candidatura. (2007). Dossier di Candidatura Expo 2015
	2010 CERTeT Study	• 'L'impatto di EXPO 2015 nell'economia italiana'. 24 Nov. 2010, 7 p. • 'Expo Milano 2015 l'impatto sull'economia italiana'. 5 p. containing the main results and available online. • 'L'impatto di expo 2015 sull'economia italiana, I risultati dell'analisi d'impatto'. Nov. 2010. 16 p. technical memo obtained at our request.
	2013 Dell'Acqua et al for Milan chamber of commerce and Expo 2015 S.p.A.	• '*L'indotto di Expo 2015. Analisi d'impatto economico'*. 20 Dec. 2013, PowerPoint presentation. 33 *slides* with 3 containing methodological considerations. • L'indotto di Expo 2015, *Un'analisi* di impatto economico, a cura di, Dell'Acqua, Q. Morri, G, Quaini E., Milano, Oct.. 2013, 102 p. Provided at our request by the authors in Jan. 2015.
Torino 2006	2005 Unione Industriale Torino study	• Unione Industriale Torino. (2005). Valutazione degli effetti economici dei Giochi Olimpici Invernali di Torino 2006. Torino: Unione Industriale Torino • Fachin, S., & Venanzoni, G. (2002). IDEM: an Integrated Demographic and Economic Model of Italy. CONSIP S.p.A.

These various documents share some similarities. For instance they are based on IO model. There are also some differences. The 2015 bidding dossier is more openly a lobbying document made for the bid, it does not have the same scientific status as others. We posit however that including this study as well is useful.

In the next section, we will present how a checklist can help to assess or improve the validity of impact studies. This will be illustrated by various cases presented in the literature and by the Italian case studies.

2. Criteria to check the validity of Economic Impact Assessments

In this section, we present a set of criteria to assess the validity of Economic Impact Assessments. We discard extrinsic conditions (e.g. independence of the study from the organisers, or experience of the analysts or balanced records of the analysts) and limit ourselves to intrinsic conditions that fully rest on the features of the study itself without consideration of the conditions in which it has been produced. Still, we include a criterion of "transparency" that is a preliminary condition for methodological inquiry, not forgetting that it is a complex criteria that has scientific but also a deontological dimension.

We present 7 macro-criteria and a decomposition at a more detailed level. Table 2 provides a synthetic presentation of these criteria.

Table 2: Check list items

Transparent

- Sufficiently detailed methodology easily available in due time
- Calculations replicable
- Key assumptions explicit, especially regarding substitution effect relating to :
  - Private expenditures: Local visitor expenditures, time switchers, etc
  - Public expenditures: investment and running of the mega-events

Educated (considers state of the art)

- Accumulated scientific knowledge recognized
- Considered substantially, not only formally

Critical toward input data

- Costs and revenues estimates
  - ex ante cost evaluation checked against risks of increase
  - Forecast of visitors and revenues checked against risk of over-estimates
- Post event impacts rest on
  - Data independent from event organisers
  - Verifiable and replicable methodology
  - Adequate number of case studies
  - Evolution of ex post flow of tourism recognizes pessimistic evaluations provided by econometric estimates
  - Evolution of Foreign Direct Investment recognizes pessimistic evaluations provided by econometric estimates

Realistic

- Reallocation of expenditures treated differently than injection of resources
  - Public expenditure substitute for Alternative use of infrastructure expenditure reasonably considered
  - Private expenditure (locals) at least partly substitute for other local expenditures
- Completeness of impacts (as compared with SAM)
  - Comprehensive of induced effects
  - Public spending impacts
  - Other SAM related impacts
- Costs defined without significant omission (accounting for security, taxes exemptions, etc.)
- Considers whether infrastructures partly built without the event (or if realization just accelerated by the event).

Balanced

- Legacy is considered both for benefits (infrastructure, image…) and losses (debts and maintenance costs, etc)
- Same level of detail for costs and benefits.

Conceptually coherent

- Allows economic benefits reduced when costs increase
- Recognizes explicitly lack of prescriptive value of economic impact analysis
- Absence of double counting
- Proper distinction between production as added value and other possible measures of economic activity
- Leakages allowed at each stage, including first one.

Territorially and temporally coherent

Territorially coherent

Results have a clear geographical scope

Area of interest coherent with territorial level financing event (if not, implications made explicit.)

IO matrix congruent with the territorial level where applied

Temporally consistent

Indication on how the age of the matrix can impact the results

Claims on temporal distribution of IO impacts are motivated

In the next section, we review one by one each of these criteria.

2.1 Transparent

A first criterion relates to how much a study is transparent. This can be assessed by investigating if a documentation is available, if it is sufficiently informative to ensure replicability and how much it makes its assumption explicit.

2.1.1 Documented

A basic criteria is that **a documentation is available**. This looks like a yes/no criteria, but our Italian case studies illustrate it is not so. For instance dell'Acqua (2013) is available months after the results were published in the newspaper. In such situation this criterion is partly fulfilled.

2.1.2 Replicable

This also means that the documentation should be sufficiently informative. A criterion is that it allows **replication**, in the meaning that an independent reader should be able to recompute the results of the study based on this document. Replicability and its companion concept of reproducibility are important in contemporary sciences, especially in psychology, medicine (Goodman et al., 2016) and, to a lesser extent, economics (Christensen & Miguel, 2018). In most IO studies, this requirement would not impose a high constraints on the users: the needed information mostly consists of an IO matrix and a demand shock vector. In other cases, the replicability condition may be quite ambitious, for instance when one uses Computable General Equilibrium (Giesecke & Madden, 2007). Even in such situations, it is, however possible and necessary that the analyst provides a full description of the model even as working notes or better sharing the model code, as advocated in computer science (Rougier et al., 2017). We are not aware of situations where this latter has ever happened in economic impact studies.

In the Italian cases we analysed, we found that replicability is not readily warranted. Sometimes minor details create indeterminacy that hurdles replication. CERTeT (2010) indicates that it uses a "2005 IO table for Italy", yet several such tables are available differing by degree of sectoral decomposition. This lack of information makes replicability harder, although analysists could proceed through trials. In such situation

gradation is appropriate to represent situations where replicability is impossible due to some indeterminacy. There is room for discretion in this rating.

2.1.3 Explicit about assumptions

Third, the **analysis should make key assumptions explicit**. This point is different from replicability: if a demand shock vector is fully documented this does not tell why it beholds given values. The assumptions that are embodied in these figures need to be accessible to the reader.

In mega-event, assumptions dealing, typically, with substitution effects of private (how much of local visitors consumption is additional, rather than substituted to other consumptions) and public expenditures are crucial Notwithstanding that the public vs. private nature of certain expenditures is partly a matter of opinion, assumptions on their additional vs. substitutive nature can change the result by an order of magnitude (Baade et al., 2005; Matheson, 2008a). Therefore, it is necessary that studies make explicit how they dealt with such fundamental assumptions.

In many cases, the assumption is visible (everything is made as if the public expenditures was fully considered additional) yet it is not explicitly stated. This is the case for two Italian studies (CERTeT (2010) and dell'Acqua (2013)). They cannot be fully compliant neither fully uncompliant if the assumption is not explicit but is clearly visible. How to assess such a situation requires to use intermediate levels in the evaluation scale. This latest point is a general indication that we guess is unnecessary to repeat for each criteria.

2.2 Critical towards input data

Another set of criteria relates to the critical use of input data.

2.2.1 Cost and revenue estimates

This applies especially to ***ex ante*** **underestimate of costs and overestimates of revenues**[2] (Althues & Maier, 2002).

For costs, the issue is both conceptual, with many 'hidden costs' not properly accounted for in many analysis (this aspect is more adequately included in the 'conceptually consistent' criteria below) but also practical, with recurrent errors in cost estimates of infrastructures and operating costs of given items. Andreff (2012) compares *ex ante* and *ex post* costs of staging the games, pointing out a staggering increase that exceeds 100% in several cases (Moscow 1980, Calgary 1988, Albertville 1992). Flyvjberg and COWI propose an upward correction of initial costs (Flyvbjerg et al., 2004).

The other side of the coin **relates to revenue estimates**. Revenues derive from fares and patronage assumptions. Arguably, the accuracy of visitors' revenues depend on the type of event considered. International exhibitions often show strong overestimates of visitors, while sport events can more often rely on predictable frequentation patterns.

To summarize, the economic impact assessment should not be naïve: costs assumptions have to be checked against risks of underestimate, and income assumption against those of overestimates. In none of the Italian studies we investigated we find explicit consideration of these issues. The question is tricky. Actually, if costs are underestimated, and substitution effects are not considered this will underestimate the impact, this open some room for opportunistic computation as underlined by Brunet (1995). Yet if substitutions effects are correctly considered this could overestimate the impacts.

2.2.2 Reasonableness of legacy effects

Critical use of data also deals with a **proper evaluation of *ex post* effects**. It is actually easy to loose contact with reality when considering all the potential benefits of an event and the lists of such benefits is potentially infinite. Actually, it is probably legitimate that analysts consider a wide array of potential effects, it is however problematic when such effects are not considered based on rigorous assessment. Ideally, analysts should rely on independent data, on a verifiable and replicable methodology and on a sufficiently large and non-distorted selection of cases. This relates, in particular, to:

- Data independent from event organisers
- Verifiable and replicable methodology.
- Adequate number of case studies
- Post event tourism flow: evidence based research has put into light that events rarely have the effects generally expected (Fourie & Santana-Gallego, 2011);
- Foreign Direct Investment: here again research indicates very limited outcomes : '*staging the Olympics has virtually no effect on FDI inflows, whereas hosting a major, nationwide football tournament might have a small positive impact on foreign investment, particularly in the years leading up to the event*.' (Jakobsen et al., 2013, p. 2).

Looking at the Italian case studies, we observe in (CERTeT, 2010) an estimate of ex post touristic impact that only refers to the case of Turin (p. 10-11). The impact on congress tourism is based only on Barcelona and, very marginally, Sydney. We suggest that in such situation the criteria of adequate number of case studies is far from being satisfactory. In saying that we do not claim that there is an easy way to deal with such issues, especially congress tourism where reliable data are scant.

2.3 Educated (considers state of the art)

Estimates should be educated, in the meaning that pre-existing research should be considered substantially.
Reference to existing literature would rather be present. Showing up with a large bibliography is an easy task, but omitting to provide references to existing work is not a valid alternative. Factually, the analysis should reference its sources and **consider**

existing scientific results. Among the Italian case studies, the CERTeT (2010) study contains no reference except to a 'recent study': *Home to big ideas: The Impact of Major Events on Inward Investment A Report by Greg Clark for Invest Thames Gateway*, but this study has been made for a group of real estate investors and cannot be classified as scientific publication. This certainly is non-compliant from this point of view.

This requirement is not formal but rather substantial. Contents of the literature ought to be considered for their stakes. If the existing literature converges on the relevance of substitution effects, it is not sufficient to include related references in bibliography. It is instead necessary **to consider substantially how such knowledge** should impact the computation.

In our set of Italian studies, Dell'Acqua et al. (2013) has a long bibliography, yet the critics formulated by the quoted literature is barely present. When present, it is disconnected from the reasoning: for instance, Baade and Matheson works (quoted as 'Baarde and Matheson') lose all their critical strength linked with the requirement of considering substitution effects. Substantial consideration of the literature is difficult to assess, but we made a deeper analysis where we looked at a series of references quoted in Dell'Acqua and checked whether their main findings were considered substantially. This analysis is available (in Italian) as supporting material. This is something that practitioners can do formally themselves with potential benefit for the quality of evaluations.

2.4 Conceptually Coherent

Being conceptually coherent is a basic requisite for any economic analysis. Applied to economic impact assessment it materializes in a few critical criteria.

A first condition for consistency relates to the relationship between **costs and the result of the Economic Impact Analysis**. In many improper economic impact analyses, especially the ones that do not adequately consider substitution, the main variable of interests is a positive function of expenditures. However, if a given infrastructure is obtained for a larger cost, it may be beneficial, for the company in charge of its construction, depending on some specific contractual agreement, but it is not beneficial for society as a whole. The resources used to cover extra costs are diverted from other socially desirable uses of public expenditures, so that the impact of these extra costs on economic activity is not necessarily positive and some loss incurs because these resources cannot be employed for other uses. A basic requirement for a proper Economic Impact Assessment is that the variables of interest are not an increasing function of the considered costs.

A second aspect is that **Economic Impact Assessment does not have a prescriptive value and should not pretend it has one**. A positive economic impact does not mean the event is 'good' for the economy '*they [IO models] attempt to measure changes in output, not welfare*' (Abelson, 2011). In order to show that an event is, in some meaning, 'good' for the economy one would need to take into consideration the opportunity costs of resources and the event's externality. Absent such considerations, the increased added value shown by many Economic Impact Assessments is not

adequate to deliver prescriptive conclusion such as 'it is good for the economy'. Normative instruments like CBA could show instead a different picture (McHugh, 2006). This limitation is unescapable when dealing with Economic Impact Assessment, failing to explicit it is improper.

A third criterion relates to **the absence of double counting**. Double counting can take an infinity of forms. We are aware of one economic impact study that examines how an increase in GDP influences the creation of new business units. The creation of extra business units is then used as an input to calculate extra added value. But there is little doubt that this just adds to the economic activity something that was already accounted for. There is also a more general issue, when the vector shock contains both the visitors' expenditures and event operating costs. The problem is that operational expenditures are partly financed by visitors spending. This implies that some impacts are accounted for twice. Checking for the lack of double counting is a daunting task: double counting comes in many various forms and can even take the form of partial double counting. While this can be object of deeper examination, we let this to future work but assume it is wise to include this criterion among our checklist.

A forth criterion relates to the clear differentiation **between production (added value) and other measures** (typically: sales). Sales are not a valid measure of economic activity. We are aware of various studies where the presented results confuse sales and value added. For instance, some of them use the word 'production' for something that is roughly the double of 'added value', without specifying what they call 'production'.

A fifth condition is that **leakages are allowed at each stage, including the first one**. This refers for instance to tourists expenditures. Some tourists my buy a tourism package abroad. Part of this expenditure will be captured by agents abroad so that they won't activate additional production in the local economy.

2.5 Realistic

Another criterion **is realism.** With due consideration to the fact that a model is always a simplification of reality**:** do the economic mechanisms present in the model exist in reality? Vice-versa, does the model include or omit effects that are relevant in the real world? This criterion is not always easy to distinguish from other categories and we recognize that certain of its items could be distributed in other categories. Yet substance should prevail on categorization. Without claim for a perfect taxonomy, this category gathers the following sub-criteria.

2.5.1 Substitution effects

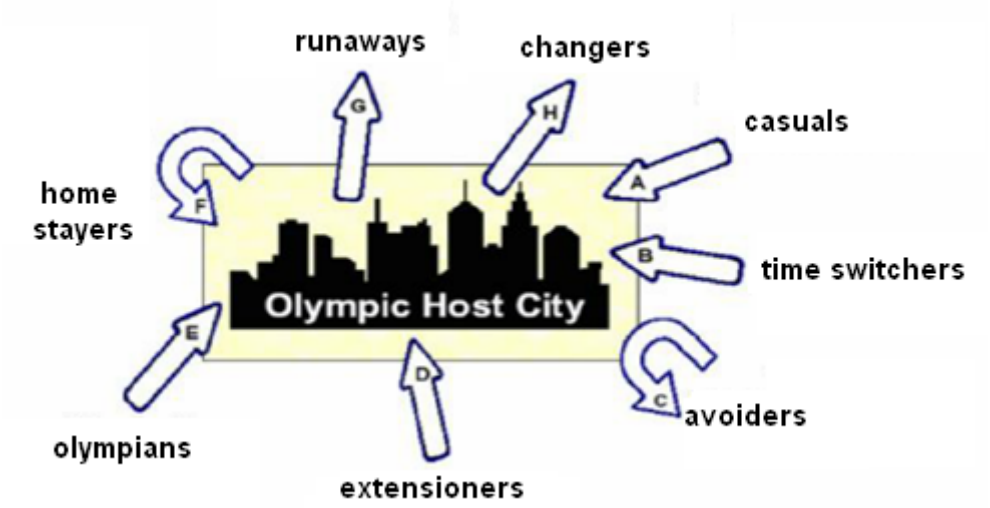
Local expenditures made for the event should not appear ex nihilo: had the event not taken place, it would have been used in other ways. This deals with public as well as private expenditure. In our terminology, reallocation of resources (for instance when public expenditures crowds out other budget out, or reduces private consumption through taxation) should be considered differently to injection.

As far **as public expenditure is concerned**, the opportunity cost of expenditure (or any equivalent concept) should be considered to reflect the cost of the foregone use of

public expenditure. For example, a study of London 2012 Summer Olympics honestly recognizes that '*the results do not consider the opportunity cost of public funds, which could be used to finance other projects or lower the tax burden.'* The report continues: *'Whilst this is a common criticism of economic impact analysis, speculating on what the funds could have been used for involves conjecture'* (Oxford Economics, 2012, p. 8). We argue that this *mea culpa* is one of the most pronounced expressions of awareness for this issue. Our analysis of Milan 2015 shows that such effects are usually omitted, they are not even mentioned: none of the studies expresses interrogations on the origin (and alternative use) of the funds used, while these questions receive increasing attention from economists.

As far **as private expenditure is concerned**, the expenditure of locals should not be considered, at least entirely, as additional: it represents an alternative use of money inside the analysed system. Nevertheless, *'[this is] frequently ignored because when expenditures by local residents are omitted, the economic impact numbers become unacceptably small to those commissioning the assessments*' (Crompton, 2006). Elaborating on this, various situations have been classified based, for instance, on the useful distinction of Preuss (Figure 1): '*Visitor expenditures should be net of 'time-switchers' and 'casuals'* (Baade et al., 2005), as already identified in the early 90's (Baade & Dye, 1990). Preuss identifies several behavioural responses to mega events. For instance, *time-switchers*, ('*visitors who had been planning to visit the study area but changed the timing of their visit to attend the event'*), *casuals*, *extensioners*, *Olympians* and *home stayers* should be considered as additional. Instead, *avoiders*, *runaways* and *changers* should not as they correspond to an alternative use of money inside the investigated area.

Figure 1: Visitor behaviours during Olympic Games



Source: Adapted from Preuss, H. (2005). The economic impact of visitors at major multi-sport events. European Sport Management Quarterly, 5(3), 281-301.

Yet, turning back to the Italian case studies, we find that none of the investigated studies actually consider this issue. They all consider event related demand by Italian households to be additional. No consideration is also made for the funding of public expenditures.

2.5.2 Consequentialism

Another aspect of realism relates to how much a given infrastructure **can really be considered a consequence of the event**. Would the city have built a subway line had the event not taken place? Or, was its construction just accelerated? If so, the vector of additional expenditures should not consider all infrastructures but only an acceleration of some of them – a tough task performed in certain CGE models (Borowski et al., 2011; for a review see: Massiani, 2018). Notwithstanding these difficulties, the main point is that full attribution of the infrastructure to the event, especially when coupled with lacking consideration of substitution effects, can cause large overestimates in economic impact assessment.

Considering our Italian case studies, they all consider a list of infrastructures provided by a central administration. Yet there is no reason to believe that all listed infrastructures would not exist without the event: a good example is Milan subway line 5 which had already been decided, financed and whose construction had started already at time of Milan candidature: no observer of the Milan situation would say it was a consequence of the Expo, yet it is present in various listings of infrastructures related to the Expo. In one case (Comitato di candidatura, 2006), the study states that such infrastructures would have been built in any case, yet they are accounted for in the demand shock.

2.5.3 Comprehensiveness (SAM completeness)

Another aspect of realism relates to the inclusion of various economic impacts, like induced effect and public expenditure effects. These can be labelled as Social Accounting Matrix completeness as such flows would materialize in various rows of a SAM. Apart from Intermediate Consumptions rows, this deals with net-of-tax labour revenues, net-of-tax capital revenues and taxes. The consideration of these various economic flows gives rise to various multiplier definitions (Batey, 1985; Miller & Blair, 2009). A proper SAM based multiplier would consider all these effects by construction (Emonts-Holley et al., 2015). But in a large number of Input-Output studies, only intersectoral effects are considered. This sometimes derives from the fact that a table of intersectoral flows is more readily available than the wider representation of the economy necessary to convey other effects. In the real world, sales finance wages (and final consumption) and government budgets (and expenditures). A computation that would fail to consider these effects would be distorted and would underestimate the event impact.

This criterion can be decomposed for convenience into:

- Indirect effects,
- Induced effects,
- Other effects, linked to other rows of the SAM, typically capital remuneration.

By construction, a proper SAM multiplier study would consider all these impacts.

2.5.4 Cost coverage

Another condition for realism relates to the **sufficient coverage of costs used in the analysis**. This question is different from the risk of cost overruns, and is more conceptual: it deals with some cost categories that may have been forgotten. There is sometimes a confusion between the (operating and investing) budget of the organizing committee or any specific body in charge of the event) and the total cost of the event, which entails a number of other expenditures like, for instance, security (Baade et al., 2005) and tax exemptions. It is then necessary that a number of auxiliary costs are properly included in the analysis.
Yet this proper consideration may sometimes create inconsistencies: in many incoherent applications of IO computation, the higher the costs, the better the outcome. This thus underlines how the audit of a given model has to consider all aspects together, an additional reason we posit for a checklist to be useful.

2.6 Balanced

Another criterion relates **to the balanced treatment of costs and benefits**. For instance, considering infrastructure heritage or image benefits is legitimate but, on the other side, one should consider interest on debts, maintenance costs and other possible deadweight of the event. We see here two subcriteria:

- The first corresponds to the consideration of both positive and negative aspects of legacy.
- The second relates to the same level of detail in costs and benefits. It is possible that the more accurately one decomposes and analyses a phenomenon, the higher will be its estimated magnitude, so that a more detailed estimate of benefits could distort the result.

This criterion is difficult to control as the number of possible impacts is very high. The Italian case studies that we analysed suggest however that it is possible to form an educated judgement on this issue. For instance, none of the studies considers the financial or the deadweight cost of funding. Neither crowding-out nor avoiders are considered. Indeed none of them considers a possible negative impact of the event. This issues a warning.

2.7 Territorially and temporally coherent

Other criteria relate to the spatial and temporal consistency of the analysis.

2.7.1 Territorial consistency

First, results should have **a clear geographical scope.** In many cases, the geographical scale may be evident, in other cases not. If so the implication of this choice should be discussed. This seems an obvious requirement, but some studies fail to do so. Crompton (1995) quotes an analysis of the Victoria 1994 Commonwealth Games:

A major problem with the study is that it provides no formal definition of the region on whose economy the impact of the Games is supposed to occur. Is it the City of Victoria or the Province of British Columbia? The study appeared to measure visitors with respect to Victoria, thus counting the residents of British Columbia from outside of Victoria as visitors. [The study appeared to measure visitors with respect to Victoria, thus counting the residents of British Columbia from outside of Victoria as visitors. […] the consultants were defining visitors with respect to the city. Yet they measured economic impact […] on the province of British Columbia, in which Victoria is a small part.] (Centre for South Australian Economic Studies, 1992, p. 11).

In the Italian cases that we analysed, we found that various studies were navigating among levels, mostly national and regional. Some ambiguities appeared in the results whose understanding probably left too much to the reader. An exception was the Turin study that was unambiguous on the geographical scale of the results.

Apart from explicating the area of interest, another point deals with the **consistency between the area of interest and the jurisdiction financing the event**. Consider an International Exhibition where the nation is ultimately providing financial guarantee to the licensing organisation. In this case, an analysis that only computes the regional effects will tell a too limited part of the story: it could be that the region economy will benefit from the event, but the main question is whether it will generate net costs or net benefits for the larger community financing the event. This issue was very present, for instance, in the 1992 Barcelona Olympics, with important transfers from the central government to Catalogne region (Solé Tura & Subirats, 1994). Obviously, providing 'regional' results is informative, but **the limitations resulting from such a choice should be explicit**. Among the 4 studies we considered in Italy, three entail national results and thus fit with this criterion: only the candidature dossier (Comitato di candidatura, 2006) is not precise on the scale of impacts and failing to discuss effectively the issue.

Another issue of spatial consistency relates to the ***ex abrupto*** **transfer of multipliers** from one territorial scale to another. The most discussible choice relates to the use of technical coefficients defined at the national level to estimate regional impacts. It should be obvious that technical coefficients are contingent to a given territorial scale. The risk otherwise is to overestimate impacts.
In the Italian studies we investigated, we found one study (CERTeT, 2010) that refers only to a national matrix but estimates regional impacts. Unless this is a documenting issue, this is a flaw.

2.7.2 Temporal consistency

Space is not everything, time matters as well. This has two particular implications. The first one is that the available IO matrix cannot be perfectly updated. This is even truer when one considers ex ante study that may be achieved 8-10 years before the event and usually use matrices that are already a few years old. This situation is unescapable and updating matrices would only be a partial solution. Analysts should however reflect about **the implications for their results**, provide the readers with fair information on this limitation and discuss the benefits and feasibility of matrix update procedures as proposed in the literature (Jackson & Murray, 2004; Parikh, 1979; Valderas-Jaramillo et al., 2019; Wang et al., 2015). In the Italian cases we analysed, this is broadly overlooked and the information provided summarizes mostly in the year of the IO matrix, with no discussion on the possibility and implications of using an outdated matrix or to update it.
The second aspects relate to the mostly **atemporal nature of the IO model**. It is conventional and mathematically consistent to see IO multiplier as accounting for the effects of a given shock until infinity (the multiplier is obtained by summing a mathematical series up to infinity). Realistically there is no reason to believe that the multiplying effect will exhaust within a limited time. This works backward as well: to produce cars you need steel to be produced before, and ore extracted even before. In all cases, unless additional assumptions are considered, it is at least discussible **to make claims about the time distribution of the benefits**. For instance dell'Acqua (2013) provides a distribution of event impacts by period. But it seems to confuse the distribution of shock with the distribution of impacts: if a demand shock takes place in a given period, this does not imply that its impact fully takes place in the same period. If this is an assumption this should at least be made explicit.

With this latter criterion we have concluded the presentation of our checklist. Yet we are aware that this may have raised additional questions. Do the application of these criteria leave too much to subjectivity? Can this checklist be used to assess works made by others or is it mostly usefull for self assessment? What would look like the full picture of the Italian case studies that we discussed?

3. Benefits and limits of a check list: an illustration based on Italian case studies

With these criteria defined, a proper analysis of the validity of events impact studies can be performed. We recognize that some of these criteria may not be straightforward. In this section, we discuss the merits and possible limits of the proposed approach. Then, we summarize the results of our Italian case study to exemplify the method.

3.1 Discussion on the methodology

Several topics are relevant. A first one is whether this check is mostly directed to self-assessment or is suitable for evaluating the quality of third part studies. We see no reason to exclude this latest, although it may make it a more controversial tool. Some other topics deserve deeper investigation. Three topics seem relevant: the question of how such a check list could be "objective", the question of assessing the global quality of a study using multivariate ordinal scales, and the question of criteria that cannot be assessed because information is missing.

3.1.1 The issue of subjectivity

Some readers of a preliminary version of our analysis pointed out that the implementation of some criteria left room to subjectivity. This deserves consideration, the question of subjectivity is pervasive in sciences especially in social sciences (Seidman, 1994) although the interest in economics is occasional (Caldas, 2016). Interestingly, economists often reflect on subjectivity referring to heterogeneity of individual tastes and less often on the value judgements that they themselves convey in their analysis.

In contrast, the notion of objectivity suggests the idea that the same tasks performed by various persons should provide the same results. This suggests the idea of a testable, deterministic or reproducible outcome. Everything that would not fit these criteria would be subjective. In its more radical form, this implies that an informed non testable statement has no better value than an uninformed non testable statement.

How would this apply to the proposed check-list? First, a number of criteria are factual, they are prone to be checked by a testable, deterministic, reproducible procedure. These refers to criteria such as:

- use of matrix at a sub-territorial level (e.g. national for regional impacts),
- age of the matrix,
- non consideration for cost overruns,
- temporal consistency.

On the other side, other criteria may not be so deterministic in their application. In many cases, the degree of conformity will be intermediate, and the various analysts could choose a different rating. The question is then whether, in such circumstances, it is beneficial for the progress of knowledge and the improvement of practice to have such criteria to check economic impact claims.

To answer this question, one should consider that we actually deal with meta-statements: Statements on Economic Impact Statements. If Economic Impact Statements would be testable and reproducible, they would generally leave little room to non-reproducible meta-statements on their validity. But economic impact statements are subjective statements. This may be unclear even to practitioners in a first instance who may be reinsured thinking they just apply a mathematical formula. But the construction of the shock vector leaves room to many degrees of freedom and

personal judgement. Not to mention all issues raised by the selection of the matrix of technical coefficients. This implies that Economic Impact Assessment are "subjective". Then the choice is between:

1. Accepting subjective claims of EIA, rejecting checking criteria because of their subjectivity.
2. Accepting such EIA assessments, whatever subjectivity they entail, so that the subjectivity present in EIA is more informed and reflexive.

Restricting the checks to criteria that are prone to determinist assessment will reduce the effectiveness of the checks, and will be more tolerant to misconduct or methodological flaws. For instance, in our Italian case studies, a report states: "*We considered a growth rate directly related to the event of 10% per year, compatible with what happened in cities who hosted international events and with congress industry similar to the one in Milan*"[3] which finally roughly transfers results from a single case (Barcelona) without quoting sources of the data. It is fair to say that the quoted example does not correspond to a fully objective calculation and it leaves large room to the analyst (how many and which reference cases will be chosen, how the differences between the reference case and the application case will be considered). It would then be problematic to accept similar subjectivity in the analysis and to exclude 'discretionary' evaluation criteria.

An interesting research would be to provide the same list of criteria to a set of reviewers and to compare their evaluations of a single study. Such a study would yet require considerable time and resources and cannot be performed at this stage of our research. We hope it will be in the near future.

We thus conclude that the proposed check list leaves some room for subjectivity on some criteria but that it is beneficial for assessing economic impact claims that are in turn highly subjective.

3.1.2 Quantifying quality with ordinal scales

The next question is whether we can derive scores or synthetic ratings of a given study and if so, compare the quality of various studies. The proposed scale is ordinal, so that assigning a quantity to each level cannot be interpreted as a quantitative scale. Such comparison would require that a weight is assigned to each criterion and that the difference between successive levels be quantified as well.
Such aggregation yet is not uncommon, for instance in multi-criteria analysis where analysts frequently aggregate several ordinal variables, often together with cardinal variables, in a general score (Hinloopen et al., 2004). Some aggregation procedures are proposed in the medical (Wittkowski et al., 2004) and psychometric literature, but they mostly rely on some common outcome variable. A possibility would be to derive the necessary weights from a pool of experts, through some preference elicitation process (Giove et al., 2008; Troutt et al., 1997). An interesting prospect would be to use Stated Preferences surveys to derive trade-off among criteria levels: it would mean for

instance that one could compare the relative weight of being non-compliant for criteria "Cost estimates checked against risks of overruns" with being uncompliant in "using an IO matrix at the improper territorial scale". This would project the various items and their various levels on a single scale. This goes beyond the scope of the present article but deserves to be discussed among economists

Another possibility would be to formulate assumptions on the statistical distribution of the underlying feature of each criteria and to show that under certain conditions a larger score implies a larger probability of performing better. The literature on quality control is going a bit further on this road using ordinal proximity measures (Franceschini et al., 2004; García-Lapresta & González del Pozo, 2019; González del Pozo et al., 2020). This prospect goes beyond the scope of the present paper and suggests further researches.

Unless we proceed with one of the above extensions, score based on the ordinal scale can have a heuristic value, so as the comparison of scores of different studies.

3.1.3 When information is missing on a criteria

Eventually, there is room for discussion about the correct way to evaluate criteria when no sufficient information is made available by the author of the study. On the one side, one should consider that failing to provide sufficient information is not similar to failure to respect a criterion. But, on the other side, a precautionary principle and a fair recognition of good documentation practice, is to consider that results that do not provide certainty on the application of the criteria should be treated like inappropriate. How to deal with such situation probably deserves discussion among economists, and can better be assessed looking at practical applications of the checklist.

This is what we do in the next section of this paper looking at Milan 2015 impact studies.

3.2 Italian case study results

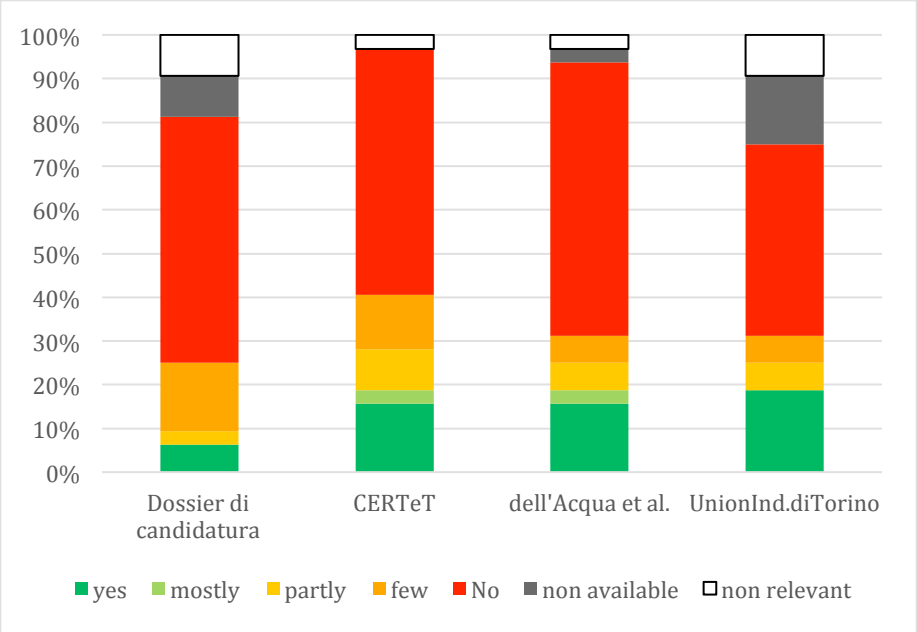
In this section, we provide an example of application of the chekclist on four studies. We see two reasons for presenting this result. First, it provides the reader a complete illustration of the possible results of a checklist. Second it provides a snapshot on a real world major event that can be informative for the community of tourism economists or for experts interested in the specific impact of Milan 2015 expo.

We make use of a five levels verbal evaluation (no, a few, partly, mostly, yes) that could easily be converted, for representation purpose, in a 5 levels ordinal scale. In many cases it is easy to make a binary yes/no assessment but, as part 2 has illustrated, this is not always the case. Intermediate situations emerge.

Sometimes a criterion is not applicable to a given study. For instance Dell'Acqua et al. (2013) study explicitly excludes infrastructure expenditures, so it is meaningless to check whether substitution effects in infrastructure spending is respected. This is denoted as Not Relevant (NR). A more critical situation is when the available documentation does not allow verification of a given criterion. A pragmatic solution is

to denote such situations with NA (Non Available). There is room for discussion about such situations. Consider two studies. One honestly recognizes a limitation in its methodology (e.g. not considering leakage in the first stage of expenditures). Another one does not say anything on this; did it just use another assumption or did it just fail to state the same assumption (one may not explicit it considers leakages in the first stage of expenditures, if he thinks it is just the sensible way to do things)? A prudential approach would dictate that the lack of certainty for the respect of a given criteria should be treated as non-conformity: as long as one is not sure that a criterion has been respected one should treat it as if it was wrong. This appears advisable although we reckon that, in some occurrences, it makes the assessment more severe.
We present hereafter the summary results and subsequently show how the evaluation can vary for single criteria. The table presenting the full results is presented in appendix. Globally, the four studies perform low in respecting the proposed criteria (Figure 2). So that this very low score are only informative of considerable issues in the practice of Economic Impact Analysis[4].

Figure 2: Compliance of 4 studies with our proposed criteria



n.a. : non available; n.r. : not relevant.

Considering single criterion, the highest level of correspondence is for “territorially coherent”, although the average level is below “partly”, the midpoint of the proposed scale. On the other side, some criteria are strikingly overlooked by all 4 studies. For all 4 studies, the following criteria have the lowest or close to lowest level:

- Considering cost overruns in Public expenditures
- Initial cost evaluation checked against risks of increase
- Substitution:

    - Public expenditure substitute for alternative use of infrastructure expenditure reasonably considered
    - Private expenditure (locals) at least partly substitute for other local expenditures
- Same level of detail for costs and benefits
- Economic benefits reduced when costs increase
- Recognizes explicitly lack of prescriptive value of economic impact analysis
- Costs defined without significant omission (accounting for security, taxes exemptions, etc.)

The list becomes longer if one also deals “n.a.”, non-availability of information to check for the respect of the criteria, as a non-conformity.

Conclusion

In this paper, we have reviewed how the definition of evaluation criteria can assist analysts or policy advisors in evaluating the adequacy of a given economic impact study. This methodology elaborates on the benefits associated with checklist and reasonableness checking (medicine, engineering, etc) that appear beneficial as long as they are conceived as a complement rather than as a substitute for reflection. These checking are mostly suitable for IO based IEA and their generalization to SAM multipliers.

Based on the evidence gathered by economic analysis in the last decades, we have formalized key requirements relating to the analysis being:

1. Transparent
2. Conceptually coherent
3. Educated (state of the art)
4. Critical in the use of data
5. Realistic
6. Balanced
7. Territorially and temporally consistent

We provide a further decomposition into sub criteria and 32 elementary criteria.

In an application to four studies made for latest mega events in Italy, the checklist appears as a valid tool to make apparent some limitations of studies. This application makes apparent that some discretion is exerted by the analyst so that a naïve claim for objectivity would not be convincing. When we shared the content of this paper with other fellows and with students we collected comments that brought us to change some evaluations given in draft versions of this paper. This indicates that the analyst may need some time to stabilize its assessment of a given study. Yet we indicate that if subjectivity should be eliminated from impact studies field, the first consequence would be to eliminate Economic Impact Assessment themselves in that they massively owe to analysts’ judgements, although the analyst is not fully aware of this. In this

setting, it is fair that an informed yet contingent effort to check for the reasonableness can improve the practice of Economic Impact Assessment.

The proposed tool can give rise to future developments following various directions. First, a more consolidated definition of criteria could be searched. The definition of the various criteria, and their possible levels, could be based on expert consultation. Another extension would be to assign weights to the various criteria and their levels. This aggregation of criteria as well could benefit from consultation of experts for instance through a Stated Preferences experiment. An important issue about aggregation is how it should allow for veto criteria, i.e. that would be sufficient to invalidate a study (but candidates for such veto would be numerous as virtually any non-conformity could dramatically reduce the validity of the study) and for non-linear effects or non-compensation. Another interesting research would be to test for the stability in the application of the criteria to a sizeable numbers of users: how would the results of the checklist vary when applied by different people.

These extensions would certainly increase the validity of the proposed methodology. We reckon however that before such extensions are made, our proposed method should be shared with the community of practitioners and scientists involved in the evaluation of mega-events.

Acknowlegment

We recognize the input of Pietro Trocciola in making this research. Although it was not possible to secure his contribution in the finalization of the present paper we acknowledge his input and thank him warmly.

References

Abelson, P. (2011). Evaluating Major Events and Avoiding the Mercantilist Fallacy. *Economic Papers: A Journal of Applied Economics and Policy*, *30*(1), 48–59. https://doi.org/10.1111/j.1759-3441.2011.00096.x

Airoldi, A., Cini, T., Morri, G., Quaini, E., & Senn, L. (2010). *L'impatto di EXPO 2015 nell'economia italiana* [CERTeT]. Università Bocconi.

Andreff, W. (2012). The winner's curse: Why is the cost of sports mega-events so often underestimated? In W. Maennig & A. (Eds.), *Zimbalist, International Handbook on the Economics of Mega Sporting Events* (pp. 37–69). Edward Elgar publishing.

Baade, R. A., Baumann, R., & Matheson, V. (2005). *Selling the Big Game: Estimating the Economic Impact of Mega-Events through Taxable Sales*.

Baade, R. A., & Dye, R. F. (1990). The Impact of Stadium and Professional Sports on Metropolitan Area Development. *Growth and Change*, *21*(2), 1–14. https://doi.org/10.1111/j.1468-2257.1990.tb00513.x

Baade, R. A., & Matheson, V. (2016). Going for the Gold: The Economics of the Olympics. *Journal of Economic Perspectives*, *30*(2), 201–218. https://doi.org/10.1257/jep.30.2.201

Batey, P. W. J. (1985). Input-output Models for Regional Demographic-Economic Analysis: Some Structural Comparisons. *Environment and Planning A: Economy and Space*, *17*(1), 73–99. https://doi.org/10.1068/a170073

Borowski, J., Boratynski, J., Czerniak, A., Dykas, P., Plich, M., Rapacki, R., & Tokarski, T. (2011). The Long-term Effect of EURO 2012 on Poland's Economy. *Ekonomista*, *4*, 493–525.

Brunet, F. (1995). An economic analysis of the Barcelona "92 Olympic Games: Resources, financing and impact." In *The Keys to success: The social, sporting, economic and communications impact of Barcelona'92*.

Caldas, J. C. (2016). *Facts, Values and Objectivity in Economics* (1 edizione). Routledge.
Centre for South Australian Economic Studies. (1992). Estimated economic impact of the 1994 Commonwealth Games. *Sport Economics*, *10-12.*

CERTeT. (2010). *Expo Milano 2015 l'impatto sull'economia italiana*.

Chang, A. S., Sin-Jing, D., & Fang-Ying, S. (2012). Engineer Self-Evaluation Checklist for Effective Site Visits. *Journal of Construction Engineering and Management*, *138*(10), 1220–1229. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000530

Christensen, G., & Miguel, E. (2018). Transparency, Reproducibility, and the Credibility of Economics Research. *Journal of Economic Literature*, *56*(3), 920–980. https://doi.org/10.1257/jel.20171350

Comitato di candidatura. (2006). *Dossier di Candidatura Expo*.

Crompton, J. L. (1995). Economic Impact Analysis of Sports Facilities and Events: Eleven Sources of Misapplication. *Journal of Sport Management*, *9*, 14–35.

Crompton, J. L. (2006). Economic Impact Studies: Instruments for Political Shenanigans? *Journal of Travel Research*, *45*(1), 67–82. https://doi.org/10.1177/0047287506288870

Davies, L., Coleman, R., & Ramchandani, G. (2013). Evaluating event economic impact: Rigour versus reality? *International Journal of Event and Festival Management*, *4*(1), 31–42. https://doi.org/10.1108/17582951311307494

Degani, A., & Wiener, E. L. (1990). *Human factors of flight-deck checklists: The normal checklist*.

Dell'Acqua, A., Morri, G., & Quaini, E. (2013). *L'indotto di Expo 2015. Analisi di impatto economico*. SDA Bocconi.

Diedering, M., & Kwiatkowski, G. (2016). Economic Impact of Events and Festivals on Host Regions—Methods in Practice & Potential Sources of Bias. *Polish Journal of Sport and Tourism*, *22*(4), 247–252. https://doi.org/10.1515/pjst-2015-0033

Drury, C. G. (2006). Human Factors and Ergonomics Audits. In *Handbook of Human Factors and Ergonomics* (pp. 1106–1132). John Wiley & Sons, Ltd. https://doi.org/10.1002/0470048204.ch42

Dwyer, L., Mellor, R., Mistilis, N., & Mules, T. (2000). A Framework for Assessing "tangible" and "intangible" Impacts of Events and Conventions. *Event Management*, *6*(3), 175–189. https://doi.org/10.0000/096020197390257

Emonts-Holley, T., Ross, A., & Swales, J. K. (2015). Type II errors in IO multipliers. In *Working Papers* (No. 1504; Working Papers). University of Strathclyde Business School, Department of Economics. https://ideas.repec.org/p/str/wpaper/1504.html

Flyvbjerg, B., Glenting, C., & Rønnest, A. (2004). *Procedures for Dealing with Optimism Bias in Transport Planning* (SSRN Scholarly Paper ID 2278346). Social Science Research Network. https://papers.ssrn.com/abstract=2278346

Fourie, J., & Santana-Gallego, M. (2011). The impact of mega-sport events on tourist arrivals. *Tourism Management*, *32*(6), 1364–1370. http://dx.doi.org/10.1016/j.tourman.2011.01.011

Franceschini, F., Galetto, M., & Varetto, M. (2004). Qualitative Ordinal Scales: The Concept of Ordinal Range. *Quality Engineering*, *16*(4), 515–524. https://doi.org/10.1081/QEN-120038013

García-Lapresta, J. L., & González del Pozo, R. (2019). An ordinal multi-criteria decision-making procedure under imprecise linguistic assessments. *European Journal of Operational Research*, *279*(1), 159–167. https://doi.org/10.1016/j.ejor.2019.05.015

Gawande, A. (2011). *Checklist Manifesto* (Reprint). Picador Paper.

Giesecke, J., & Madden, J. (2007). *The Sydney Olympics, seven Years On: An Ex-Post Dynamic CGE Assessment*. Centre of Policy studies Monash University.

Giove, S., Rosato, P., & Breil, M. (2008). A Multicriteria Approach for the Evaluation of the Sustainability of Re-use of Historic Buildings in Venice. In *Sustainability Indicators and Environmental Valuation Working Papers* (No. 46625; Sustainability Indicators and Environmental Valuation Working Papers). Fondazione Eni Enrico Mattei (FEEM). https://ideas.repec.org/p/ags/feemsi/46625.html

González del Pozo, R., Dias, L. C., & García-Lapresta, J. L. (2020). Using Different Qualitative Scales in a Multi-Criteria Decision-Making Procedure. *Mathematics*, *8*(3), 458. https://doi.org/10.3390/math8030458

Goodman, S. N., Fanelli, D., & Ioannidis, J. P. A. (2016). What does research reproducibility mean? *Science Translational Medicine*, *8*(341), 341ps12-341ps12. https://doi.org/10.1126/scitranslmed.aaf5027

Haynes, A. B., Weiser, T. G., Berry, W. R., Lipsitz, S. R., Breizat, A.-H. S., Dellinger, E. P., Herbosa, T., Joseph, S., Kibatala, P. L., Lapitan, M. C. M., Merry, A. F., Moorthy, K., Reznick, R. K., Taylor, B., & Gawande, A. A. (2009). A Surgical Safety Checklist to Reduce Morbidity and Mortality in a Global Population. *New England Journal of Medicine*, *360*(5), 491–499. https://doi.org/10.1056/NEJMsa0810119

Hinloopen, E., Nijkamp, P., & Rietveld, P. (2004). Integration of ordinal and cardinal information in multi-criteria ranking with imperfect compensation. *European Journal of Operational Research*, *158*(2), 317–338. https://doi.org/10.1016/j.ejor.2003.06.007

Hudson, I. (2001). The Use and Misuse of Economic Impact Analysis: The Case of Professional Sports. *Journal of Sport and Social Issues*, *25*(1), 20–39. https://doi.org/10.1177/0193723501251003

Jackson, R., & Murray, A. (2004). Alternative Input-Output Matrix Updating Formulations. *Economic Systems Research*, *16*(2), 135–148. https://doi.org/10.1080/0953531042000219268

Jakobsen, J., Solberg, H. A., Halvorsen, T., & Jakobsen, T. G. (2013). Fool's gold: Major sport events and foreign direct investment. *International Journal of Sport Policy and Politics*, *5*(3), 363–380. https://doi.org/10.1080/19406940.2012.717099

Jeanrenaud, C. (2007). Sports Events: Uses and Abuses of Economic Impact Studies. *Finance & Bien Commun*, *No 26*(1), 99–104. https://doi.org/10.3917/fbc.026.0099

Késenne, S. (2005). Do We Need an Economic Impact Study or a Cost-Benefit Analysis of a Sports Event? *European Sport Management Quarterly*, *5*(2), 133–142. https://doi.org/10.1080/16184740500188789

Leontief, W. (1966). *Input-output economics*. Oxford University Press.

Massiani, J. (2018). Assessing the economic impact of mega events using Computable General Equilibrium models: Promises and compromises. *Economic Modelling*, *75*, 1–9. https://doi.org/10.1016/j.econmod.2018.05.021

Matheson, V. (2008a). Mega-events: The effect of the world's biggest sporting events on local, regional, and national economies. *Humphreys, The Business of Sports*, *1*, 81–99.

Matheson, V. (2008b, February 14). Upon Further Review: An Examination of Sporting Event Economic Impact Studies. *The Sport Journal*. http://thesportjournal.org/article/upon-further-review-an-examination-of-sporting-event-economic-impact-studies/

Matheson, V. (2009). Economic Multipliers and Mega-Event Analysis. *International Journal of Sport Finance*, *4*(1), 63–70.

McHugh, D. (2006). *A Cost-Benefit Analysis of an Olympic Games*. Queen's Economics Department.

Miller, R. E., & Blair, P. D. (2009). *Input Output Analysis – Foundations and extension* (2nd edition). Prentice-Hall, Inc.

Mondello, M. J., & Rishe, P. (2004). Comparative Economic Impact Analyses: Differences Across Cities, Events, and Demographics. *Economic Development Quarterly*, *18*(4), 331–342. https://doi.org/10.1177/0891242404269505

Mules, T., & Faulkner, B. (1996). An economic perspective on special events. *Tourism Economics*, *2*(2), 107–117. https://doi.org/10.1177/135481669600200201

Nicolas, Y. (2007). *Les premiers principes de l'analyse d'impact économique local d'une activité culturelle* (Ministère de La Culture. Paris. Culture Méthodes - Production, Diffusion et Marchés, p. 8).

Oxford Economics. (2012). *The economic impact of the London 2012 Olympic & Paralympic Games. Oxford Economics.* (p. 52). commissioned by Lloyds Banking Group.

Parikh, A. (1979). Forecasts of Input-Output Matrices Using the R.A.S. Method. *The Review of Economics and Statistics*, *61*(3), 477–481. JSTOR. https://doi.org/10.2307/1926084

Porter, P. K. (1999). Mega-sports events as municipal investments: A critique of impact analysis. *Sports Economics: Current Research*, 61–73.

Preuss, H. (2005). The Economic Impact of Visitors at Major Multi-sport Events. *European Sport Management Quarterly*, *5*(3), 281–301. https://doi.org/10.1080/16184740500190710

Rougier, N. P., Hinsen, K., Alexandre, F., Arildsen, T., Barba, L. A., Benureau, F. C. Y., Brown, C. T., Buyl, P. de, Caglayan, O., Davison, A. P., Delsuc, M.-A., Detorakis, G., Diem, A. K., Drix, D., Enel, P., Girard, B., Guest, O., Hall, M. G., Henriques, R. N., … Zito, T. (2017). Sustainable computational science: The ReScience initiative. *PeerJ Computer Science*, *3*, e142. https://doi.org/10.7717/peerj-cs.142

Seidman, S. (1994). *The Postmodern Turn: New Perspectives on Social Theory*. Cambridge University Press.

Siegfried, J. J., & Zimbalist, A. (2000). The Economics of Sports Facilities and Their Communities. *Journal of Economic Perspectives*, 95–114.

Solé Tura, J., & Subirats, J. (1994). *La organización de los Juegos Olímpicos de Barcelona'92: Un ejemplo de economía mixta o de sociedad pública-privada.*

TMIP. (2010). *Travel Model Validation and Reasonability Checking Manual* (Second Edition). Transport Modelling Improvement Program. Federal Highway Administration.

Troutt, M. D., Rai, A., & Tadisina, S. K. (1997). Aggregating multiple expert data for linear case valuation models using the MDE principle. *Decision Support Systems*, *21*(2), 75–82. https://doi.org/10.1016/S0167-9236(97)00018-3

Tyrrell, T. J., & Johnston, R. J. (2001). "A Framework for Assessing Direct Economic Impacts of Tourist Events: Distinguishing Origins, Destinations, and Causes of Expenditures. *Journal of Travel Research*, 94–100.

Valderas-Jaramillo, J. M., Rueda-Cantuche, J. M., Olmedo, E., & Beutel, J. (2019). Projecting supply and use tables: New variants and fair comparisons. *Economic Systems Research*, *31*(3), 423–444. https://doi.org/10.1080/09535314.2018.1545221

Wang, H., Wang, C., Zheng, H., Feng, H., Guan, R., & Long, W. (2015). Updating Input–Output Tables with Benchmark Table Series. *Economic Systems Research*, *27*(3), 287–305. https://doi.org/10.1080/09535314.2015.1053846

Wittkowski, K. M., Lee, E., Nussbaum, R., Chamian, F. N., & Krueger, J. G. (2004). Combining several ordinal measures in clinical studies. *Statistics in Medicine*, *23*(10), 1579–1592. https://doi.org/10.1002/sim.1778

Wyk, L. V., Saayman, M., Rossouw, R., & Saayman, A. (2015). Regional economic impacts of events: A comparison of methods. *South African Journal of Economic and Management Sciences*, *18*(2), 155–176. https://doi.org/10.4102/sajems.v18i2.593

Zimbalist, A. (2010). Is It Worth it? Hosting the Olympic Games and other mega sporting events is an honor many countries aspire to—But why? *Finance & Development*, 8–11.

Zimbalist, A. (2015). *Circus Maximus: The Economic Gamble Behind Hosting the Olympics and the World Cup*. Institution Press.

[1] Another existing document is: Dell'Acqua, A., Etro, L. L., (2008), 'Expo Milano 2015. Un'analisi di impatto economico per il Sistema Paese ed i settori industriali italiani', SDA Bocconi School of Management, Milan, Research report. This does not appear to be publicly available.

[2] For example, 'most league-sponsored economic impact studies (…) often completely ignore the costs of hosting such an event' (Matheson, 2008) *or* 'While on paper the 2002 Winter Olympics in Salt Lake City made a profit, the cost figures did not include millions of dollars of additional security provided by the U.S. Department of Defense at no cost to the local organizing committee' (Baade, Baumann, & Matheson, 2005)

[3] 'Si è considerato un tasso di crescita direttamente attribuibile del 10% l'anno, compatibile con quanto avvenuto in città che hanno ospitato eventi internazionali e con un mercato congressuale simile a quello Milanese' (Airoldi et al., 2010).

[4] It is tempting to compute a synthetic score. A simple (unweighted) summation (assigning values (0; 1; 2; 3; 4) to the 5 levels of the verbal scale and an equal weight for each elementary criteria of the scores of the 4 studies would produce very low validity scores. Yet we are aware this metric is imperfect, as it considers that all criteria are of equal importance, and that the difference between each level in each criteria is identical. Such result would range from 7% to 20 % of the maximum score (this maximum score excludes criteria that are not relevant for a given study), yet the validity of the computation is limited when dealing with multivariate ordinal scales. The value of such computation would mostly be heuristic.

Appendix: detailed check list for Milan 2015 and Turin 2006 economic impact studies

Evaluation levels: no, few, partly, mostly, yes. ; n.a : non available, n.r.: Not relevant

Criteria Sub criteria Elementary criteria	Milano 2015 Dossier di candidatura (2007)[1]	CERTeT (2010)	Dell'Acqua et al (2013)	TO 2006 Union Ind. di Torino (2005)
Transparent				
Sufficiently detailed methodology easily available in due time	Few	few[2]	No[3]	Yes
Calculations replicable	Few	Partly[4]	Partly[5]	Partly[6]
Key assumptions explicit, especially regarding substitution effect relating to :				
Private expenditures: Local visitor expenditures, time switchers, etc	n.r.[7]	Partly[8]	No[9]	No
Public expenditures: investment and running of the mega-events	No	No[10]	No[11]	No
Informed				
Accumulated scientific knowledge recognized	No	No[12]	Yes	No
Considered substantially, not only formally	No	n.r.	No[13]	No
Critic				
Initial cost evaluation checked against risks of increase	No	No	No	No
Forecast of visitors number based on a rigorous methodology	Partly	Partly[14]	Partly[15]	n.a.[16]
Realistic				
Reallocation of expenditures treated differently than injection of resources				
Public expenditure substitute for Alternative use of infrastructure expenditure reasonably considered[17]	No	No	No	No
Private expenditure (locals) at least partly substitute for other local expenditures	No	No	No	No
Estimate of *ex post* effect rests on:				
Data independent from event organisers	n. a.	Yes	Yes	n. a.
Verifiable and replicable methodology	No	Yes[18]	No[19]	n. a.
Adequate number of case studies	No	No[20]	No[21]	n. r.
Evolution of *ex post* flow of tourism recognizes pessimistic evaluations provided in literature	n. r.[22]	few[23]	No[24]	n. r.
Evolution of Foreign Direct Investment recognizes pessimistic evaluations provided in literature	n. r.	No[25]	No[26]	n. r.
Considers whether infrastructures partly built without the event (or if realization just accelerated[27, 28]	No[29]	No[30]	n. r.	No[31]
Balanced				
Legacy is considered both for benefits (infrastructure, image…) and losses (debts and maintenance costs, etc) [32]	No	No[33]	No[34]	n.a.
Same level of detail for costs and benefits. [35]	No	No[36]	No[37]	No
Conceptually coherent				
Allows economic benefits reduced when costs increase	No	No	No	No
Recognizes explicitly lack of prescriptive value of economic impact analysis[38]	No	No	No	No

Costs defined without significant omission (accounting for security, taxes exemptions, etc.)	No	No[39]	No[40]	No[41]
Absence of double counting	Yes	Yes	No[42]	Yes
Proper distinction between production as added value and other possible measures of economic activity	No	No[43]	No[44]	Yes
Territorially and temporally coherent				
Choice of area of analysis	No[45]	Partly[46]	Partly[47]	Yes
Choice of area explicitly discussed,	No[48]	Yes	Yes	Yes
Area of interest coherent with territorial level financing event (if not, implications made explicit.)				
IO matrix congruent with the territorial level where applied	n.a.[49]	No[50]	Yes[51]	Yes[52]
Leakages allowed at each stage, including first one.	n.a.	No[53]	Partly[54]	n.a.
Temporally consistent[55]				
Indication on how the 'age' of the matrix can impact the results	No	Few	Few	Few
Claims on temporal distribution of benefits are duly discussed.	Partly[56]	Yes[57]	No[58]	Partly[59]

[1] Non-conformity of the bidding dossier with our criteria should be considered taking into consideration the lobbyist nature of this document. In this analysis, we may therefore be less detailed about this document than about the others.

[2] A 16 p. methodological memo has been sent to us at our request. It clarifies some of the calculations.

[3] A methodology of 102 pages has been communicated at our request one year after publication of the results and after several requests from a collaborator of ours. In between times, some claims produced by the study had been widely quoted in the media.

[4] For instance, the Input-Output matrix is not discussed. The indication 'tav.I/O 2005 Italia' may not be sufficient to warrant replicability.

[5] Matrix calculation is described in details (cap. 7.3) although the matrix itself is not documented. But other parts of the computation are not sufficiently described so as to allow for replication (for instance increased entrepreneurship and creation of new companies (l'*imprenditorialità incrementale e la creazione di nuove imprese*) is based on a model that is only partly documented p. 62. The key factor of the number of additional companies relating to an increase in GDP is not provided, although the reader may try to deduce it from other information present in the dossier.

[6] Apart from the study, a 15 pages description is available. This document does not appear sufficient to warrant the replicability of computations. One could however object that the replicability condition can be verified with varying levels of requirements based on the very complexity of a given model.

[7] Bidding dossier does not consider visitors' expenditures.

[8] The study considers a percentage of presence in Lombardy or in Milan for which 'motivation of trip is Expo'. This information is provided by a poll, but the precise way by which this information has been collected, and the precise question phrasing, is not available.

[9] Dossier p. 14. No indication is presented on the subtraction of the substitute part of the local visitors.

[10] The whole public expenditure is treated as additional for the economic system considered.

[11] The public vs. private nature of certain expenditures is partly a matter of opinion. Still, the whole public expenditure is treated as additional for the various study areas considered.

[12] No reference is made to scientific publications. The Dossier refers to a 'recent study' *Clark G. 'Home to big ideas: The Impact of Major Events on Inward Investment. London 2012 and The Thames Gateway*', but this study has been made for a group of real estate investors and cannot be classified as scientific publication.

[13] The critics formulated by existing scientific literature on mega events is barely present, and, when present, it does not impact reasoning (see for instance the use of Baade and Matheson works, quoted as 'Baarde and Matheson', which lost all its critic aspects as the ones on substitution effects.

[14] The estimate of Italian visitors is based on a survey, without consideration of how interviewed people declaration should be corrected due their nature of statement. For most of the areas of origin of the visitors, the estimate is substantially based on a doubling of visits, calculation that is not supported.

[15] The study refers to a survey ('*a survey made by Bain Italy for Expo 2015 Spa'* ; '*un'indagine svolta da Bain Italia su incarico di Expo* 2015 S.p.A' p.14; 'corrected in February 2013 based on the results of a new survey made in December 2012 by Eurisko'; *'corrette nel febbraio 2013 sulla base dei risultati della nuova indagine svolta nel dicembre 2012 da Eurisko*' p. 15) does not provide sufficient details to allow for an analysis.

[16] Visitors forecast is usually easier for Olympics than for International Exhibitions.

[17] None of the studies expresses interrogations on the origin (and alternative use) of the funds used, while these questions receive increasing attention from economic analysis (Zimbalist 2015).

[18] Some features of the computation of FDI (see for instance p.12) are exposed in a very simple manner, but this can relate more closely to the real simplicity of the underlying model than to a lack of documentation.

[19] The effect on Foreign Direct Investment is based on 'historical data linked to the dynamic of FDI in Italy and in the areas of interest of the Expo' ('*benchmark storici legati alla dinamica degli IDE in Italia e nelle aree interessate da Expo nello specifico',* p. 64) but the corresponding data are not made available. Relating to touristic flows, discussed essentially in pp. 17 and 64, the study does not provide precise information on how 1 billion stays have been obtained, other than referring to a 2008 study by the same authors, which is not publicly available. The research by (Fourie & Santana-Gallego, 2011) that consolidate observations on touristic flows between 200 countries is not quoted. It would however be excessively severe to blame the author of Milan study for this omission, as the publication was very recent.

[20] The estimate of ex post touristic impact on Milan only refers to the case of Turin (p. 10-11). The impact on congress tourism is based only on Barcelona and, very marginally, Sydney.

[21] For instance, the impact on ex post tourism only refers to the Turin case (p. 17). A study of 2008 by the same authors could contain extra information but it does not seem available to the public.

[22] The discussion on post event touristic flow (p. 126) is barely substantial and does not produce a quantification, but in any case these effects are not included in the impact study.

[23] One may refer, for instance, to the discussion on the number of congress participants, a discussion that seems highly speculative 'we considered a growth rate directly related at 10% per year, compatible with what happened in cities who hosted international events with a congress market similar to the one in Milan' p. 6. In reality, the estimate is based, although with claims of being prudential, only on the case of Barcelona, whose figures are provided without quoting their sources.

[24] p. 64: the quantification of ex post flows does not appear justified 'based on the estimates contained in the analysis, such effects should generate an additional production of 1,2 billion euros'. ('*In base alle stime contenute nell'analisi tali effetti dovrebbero generare una produzione aggiuntiva di 1,2 Miliardi di euro*'). No model is presented that would produce this figure impact or at least would deduce it from other cases. Considerations provided in pages p. 82-84 do not provide further clarification.

[25] 'it is likely, as demonstrated by studies made for other similar events, that an increase in FDI takes place for a few years' ('è probabile, come dimostrano studi realizzati in previsione di eventi analoghi, che si verifichi un aumento per qualche anno dei flussi d'investimenti diretti esteri (IDE)' p. 6). Such studies, however, are not quoted. More pessimistic results obtained in other researches are not considered.

[26] The study states 'based on indications provided by some sources, a 5% increase of 'Expo induced' FDI has been estimated' ('Sulla base di indicazioni tratte da alcune fonti, si è stimato un aumento annuo degli IDE 'Expo-induced' del 5%,') but the corresponding sources are not quoted.

[27] The study should at least deal consistently with the fact that acceleration in the construction of these infrastructures implies a reduced priority for other infrastructure or an increase in taxation.

[28] One can honestly state that none of the considered studies, which deal with infrastructures effectively, considers this question. Considering Turin, it seems implicit and certain for the authors taht the expenditures for infrastructures in periods 2001-2004 and 2005-2009 would not have taken place without the Olympics.

[29] It seems, at least implicitly, that the 10,179 billion of infrastructure expenditures (called 'infrastructural investments not linked to Expo') would not have been spent at all if Milan would not have organized the Expo.

[30] The whole infrastructural expenditures are considered (p. 5)

[31] The study refers to 'construction of infrastructure for the operation of the games and the running of the Turin organizing committee' ('realizzazione delle opere per lo svolgimento dei Giochi e al funzionamento del TOROC') and also 'related infrastructures' ('opere connesse') to be constructed in the games area, and to 'accompanying infrastructure' ('opere di accompagnamento'). The implicit assumption seems to be that none of these infrastructures would have been achieved, even partially, without the Olympics. This assumption is however discussible and, in all cases, should have been discussed by the authors, considering how unlikely it was that, absent the games, no intervention on these infrastructures would have taken place.

[32] For instance crowding out is neglected only after evaluation that the impacts of such omission are less than those of other benefit considered.

[33] Reference to maintaining costs (p. 5) is misguiding: they appear as a benefit, not as a cost.

[34] Interesting the fact that Crystal Palace is quoted as a '*still tangible and visible*' ('*ancora oggi concrete e visibili*') example of International Exhibition Legacy, forgetting that it was destroyed by fire in 1936! (p. 59)

[35] For the various available studies, it would be sufficient to observe that financial costs (interests) are omitted. Moreover, none of the studies considers crowding out, although this aspect may be considered minor. More fundamental is the lack of consideration of opportunity costs of public funds.

[36] Crowding out is not considered, neither substitution effects.

[37] The study does not deal with annex infrastructures. This could lead to an underestimate of the economic impact. On the other hand, the omission of substitution effects creates an overestimate. It thus appears unclear whether the final result is over or underestimate.

[38] For instance, a Cost Benefit Analysis could generate a negative outcome event in presence of a positive economic impact (increase in added value).

[39] The chosen methodology is not discussed and compared with prescriptive methods proposed in public economy. The lack of prescriptive value of the results is not made explicit.

[40] idem

[41] idem

[42] The study compares how increased GDP affects the constitution of new companies and considers also how increased number of companies in turn increases GDP . The computation could loop for a higher number of times, but two times is already double counting.

[43] 'Produzione aggiuntiva complessivamente determinata da EXPO Milano 2015 nell'economia italiana, nel periodo 2011-2020, potrà ammontare a più di € 69 Miliardi, cui corrisponde un incremento di valore aggiunto pari a circa € 29 Miliardi.'

[44] Refer to increased production as something different from increase in added value. Context strongly suggest 'increased production' is used to label change in output.

[45] The chapter title refers to Italy and Milan. Some figures are quoted p. 116 without explicit reference to a given area. The readers may later discover (tab. 21.1) that such figures refer to Lombardy. It may just deal with clarity in communication; it does not however help to establish the credibility of the study.

[46] The study considers national economy and regional (Lombardy) economies economy, but it is sometimes uneasy to tell apart what relates to Lombardy and what relates to Italy. This may, here again, be a matter of expression. This however does not help to establish the reliability of the study.

[47] The study considers Milan Province (an administrative subdivision extending few tens of kilometres outside of Milan), Lombardy and Italy. The territorial coverage allows for various levels of analysis, it is however unclear how impact on the province level can be estimated using a regional intersectoral table (as can be deducted reading p. 99).

[48] The study does not discuss the study area but implicitly concentrates on the regional level (see p. 119 for instance).

[49] At least it is not possible to exclude the use of an unappropriated matrix. The impact is computed at the regional level but the scale of the matrix is not clarified.

[50] 'The basis for the impact analysis is the Input-Output matrix for year 2005' ('La base per l'analisi dell'impatto è la Matrice Input Output dell'Italia relativa al 2005') (l'impatto di Expo 2015 sull'economia italiana; I risultati dell'analisi d'impatto, Novembre 2010 p. 4.)

[51] At least if we consider this statement: 'indirect activation has been estimated based on the Table of Intersectorial Trade relating to the Italian economy for 2005 and to Lombardy for year 2006' ('L'attivazione indiretta è stata stimata sulla base della TEI relativa all'economia italiana al 2005 e alla tavola Lombardia 2006') p. 99.

[52] It is a proper feature of IDEM model to take into consideration, with special attention, the regional nature of the impacts.

[53] The study recognizes, genuinely, 'generally (…) it has been assumed that al the inputs in the impact vector are provided by Italian firms' ('Generalmente, (…) si è ipotizzato che tutti gli input indicati nei vettori d''impatto provengano da imprese italiane') (L'impatto di expo 2015 sull'economia italiana; I risultati dell'analisi d'impatto, Novembre 2010 p. 8.)

[54] p. 26: a small share of visitors' final consumption leaks to Switzerland in the first stage.

[55] With some goodwill, one could understand that most of the impacts occur "shortly after' the choc, but this assumption should at least be explicit and its validity should be discussed.

[56] Some contrary statements can be found: '*70,000 new jobs during the timeframe necessary to prepare for Expo Milano* 2015.' chapter 21, p 116.

[57] Although one can read '*for precautionary reasons, it has been assumed that all impacts related to Expo Milano 2015 will be exhausted in 2020*' ('*In via cautelativa, si è supposto che tali effetti imputabili a EXPO Milano 2015 si esauriscano nel 2020*'), this statement is however rather vague, and does not fully justify the assumptions made about the temporal distribution of the effects.

[58] Refer to section 3.3 of the complete report. Starting from p. 22.

[59] Refer to chapter 3.1 and 3.2 of (Unione Industriale Torino, 2005) that display a temporal distribution of effects.