# Assessing Regulators Fairly: Matching Scientific Knowledge With Impact Assessments

Fabrizio De Francesco

Paper presented at the international symposium on Regulatory Impact Analysis organised by the European Court of Auditors
Luxembourg, 23 and 24 November 2017

## Abstract

Scorecards, retrospective analysis, and regulatory performance measures have been used for assessing the quality of impact assessments. However, these methods neglect the actual scientific knowledge available at the time of conducting a specific policy appraisal. Furthermore, retrospective large-N and single case analyses are not able to capture the extent of diachronic learning related to a specific sector of regulatory reform.

By focusing on the EU railway reform, the aim of this paper is to assess the quality of economic analysis in impact assessments while considering the extent of scientific knowledge. The research design is straightforward. A review of the economic literature traces the progress in the scientific methods for evaluating the economic impact of rail liberalisation. By matching scientific knowledge with the knowledge expressed in impact assessment practices, the main hypothesis to test is whether economists' evaluation models are fully exploited by policy appraisers. The contribution of the paper is two-fold: on the one hand, it enhances the methodology for regulatory policy evaluation and, on the other hand, it contributes to the literature of the use of scientific knowledge in policy making.

# 1 Introduction

Impact assessment (IA) is a regulatory innovation (Black, Lodge, and Thatcher 2005) that has spread transnationally (De Francesco 2013). The attainment of its institutional viability is not straightforward. Contrarily to adoption that is a highly symbolic event (Goodman and Steckler 1989), the implementation of IA is complex and goes through several phases in which the relationships and relative power positions of policy actors vary (De Francesco, Radaelli, and Troeger 2012). Similarly to other innovations (Goodman and Steckler 1989; Steckler et al. 1992), IA became institutionalised as regulators follow the required evaluation standards and practices and procedural routines.

How can one assess the institutionalisation of IA? As an instance of evidence-based policymaking, the institutionalisation of IA can have two alternative purposes and functions (Head 2016). According to the advocates of significant commitment to rigorous methodologies for program evaluation, IA practices should be assessed for the production of scientific evidence necessary for enhancing the economic efficiency of their decisions. A less ambitious, but realistic approach is to consider good policymaking based on 'a range of relevant "best available" evidence' produced by professional expertise (Head 2016: 474).

Based on the latter approach, this paper puts forwards a straightforward "follow the evidence" method for retrospectively assessing the quality of scientific methods utilised in IAs. It argues that regulators and policy evaluators should be assessed according to the *utilisation* of the best available (professional and scientific) evidence and evaluation methods, rather than the evidence *produced* in order to take the best decision.

There are several modes for retrospectively assessing IAs (see for a review of international experience OECD 1999; OECD 2003; Radaelli and De Francesco 2007; De Francesco and Radaelli 2007; Radaelli and Fritsch 2011). Although the OECD's efforts (1999, 2003), there are not yet consolidated standards for evaluating the quality of IAs, since the

governments' evaluative practices concern mainly the quantification of administrative burdens (De Francesco 2013) and the "standard cost model" (Coletti 2013). As a consequence, scholars and consultants have attempted to fill this void by proposing several methods for assessing IAs.

Retrospective analysis of IAs can be classified according to three dimensions: function, procedure, and methodological accuracy of IA. Impact studies (OECD 1999: 19) or function tests (Harrington and Morgenstern 2004: 13-14; Bizer Lechner, and Führ 2010: 33) assess whether an IA has influenced the decision making and has improved the quality of legislation. Conduct studies (OECD 1999: 19) or content tests (Bizer, Lechner, and Führ 2010: 33) aim to evaluate the gap between the procedural criteria prescribed by IA guidelines and the content of regulatory analyses (Hahn et al. 1999, Ellig and McLaughlin 2011). Finally, analytical accuracy studies (OECD 1999: 19) or result tests (Bizer et al 2010: 33) focus on the quality of the economic predictions and their agreement with the actual impact of a new regulation.

Within the latter group of retrospective analysis, this paper puts forward the methodological recommendation of tracing the evolution of economists' models (for assessing the impact of liberalisation) in order to identify general patterns of the evaluation methods and practices as represented in the IA documents (on the practice of writing government documents cf. Freeman and Maybin 2011; Wesselink, Colebatch, and Pearce 2014). The proposed method is applied to 16 IAs and studies conducted between 2004 and 2017 associated with EU rail liberalisation. An in-depth analysis of two IAs complements this qualitative assessment.

Although assessment of the extent of utilization of scientific knowledge is not novel in the (regulatory) IA literature (Desmarais and Hird [2014] and Costa, Desmarais and Hird [2016] rely on bibliographic metrics), this contribution emphasises the level of authority achieved by regulators through their IA documents vis-à-vis the scientific knowledge available at the time of a regulatory proposal. The long sequence of EU railway reform and the number of IAs and evaluation studies produced

allows also to capture the extent of diachronic learning reflected in the use of (scientific) evidence across several regulatory reform packages necessary for the completion of the liberalisation (cf. Torriti 2010 that analysed only an individual EU IA on energy liberalisation).

The remainder of this paper is structured as follows. Situating the methodological contribution of this paper, the next section evaluates the advantages and disadvantages of the existing models and practices used for assessing the quality of IAs. Section 3 provides a brief overview of the EU rail liberalisation. While Section 4 ranks and traces the evolution of the scientific methods for assessing the impact of rail liberalisation, Section 5 assess whether the available scientific knowledge has been translated into IAs. Section 6 concludes by summarising the main empirical findings and proposes policy recommendations and future avenues of research.

## 2 Modes of evaluating impact assessments

Because of the difficulty in conducting impact studies on IAs (but see (Shapiro 2008; Shapiro and Morrall III 2012), the literature on the evaluation of IAs tends to focus on two main types of retrospective analysis: conduct studies scoring the extent of "compliance" with IA guidelines and accuracy studies verifying regulatory cost and benefit estimates. Both modes of evaluation have advantages and disadvantages, as the next subsections will show.

### 2.1 Scoring conducts

Since the first scorecard drawn by Hahn et al. in 1999, ranking the conduct of policy appraisers is a common evaluative standards followed by (see Cecot et al. 2008; Fritsch et al. 2013 for scorecards applied on EU IAs). Economists (Hahn et al. 1999; Hahn and Tetlock 2008; Lee and Kirkpatrick 2006) and public policy scholars (Adelle, Hertin, and Jordan 2006), think tanks (Vibert 2004; Wilkinson et al. 2004), and stakeholders (NNR 2006) rely on this tool to evaluate the state of the art of (regulatory)

IA programmes in the USA, the UK, and at the European level. The feature of this type of retrospective analysis is to relying on guidelines in order derive the standard conduct of a regulator. For instance, in his first scorecard Hahn et al (1999) relies on the OMB-OIRA guidelines on complying with the E.O. 12,866.

Scorecards are composed of a series of Yes/No questions that generate simple measures. A feature of these measures is that they can be weighted and aggregated in an overall composite indicator. This differentiates scorecards from the checklists. The latter are usually a set of single measures that are not aggregated.

IA scorecards are usually developed according to the following qualitative dimensions: quantification or monetization of regulatory costs and benefits (Vibert 2004, Torriti 2007), the consideration of sustainable development (Wilkinson et al. 2004, Adelle, Hertin, and Jordan 2006), and the consideration of several options (Renda 2006; Cecot et al. 2008). Since the sample of regulation is usually large, scorecards have never been applied to a coherent set of regulations related to the liberalisation of a specific utility market. Accordingly, it is impossible to qualify the extent of diachronic learning of the regulatory analyst in a specific regulatory reform. Furthermore, the IA scorecard has not been designed with the purpose of assessing the extent of gap between the scientific knowledge and the knowledge expressed in the IA. This goal is better achieved by the evaluation mode that verifies the estimations of regulatory costs and benefits.

### 2.2 Verifying estimates

Several scholars have compared *ex ante* estimates of costs and benefits of specific regulations with *ex post* assessments of regulatory impact. There are three sources for discrepancies between predicted and actual impacts. The first source of inaccuracy is implicit in the uncertainty over future and unpredictable changes of: i) relative prices of cost components, ii) technology and affected parties' adaptation to new regulations, and iii) the wider economic conditions. The second source of inaccuracy is related to

the lack of scientific knowledge on cause-effect relationships. The final source of inaccuracy refers to 'uncertainties associated with modelling activities, particularly in regard to any assumptions which have to be made by analysts' (OECD 1999: 39). While the first and the second types of inaccuracy are extremely difficult to lower through an IA, the latter source of inaccuracy is what matters in assessing the quality of scientific knowledge utilisation and the appropriateness of regulatory analysts. In other words, it is important to separate discrepancies arising from flawed applications of scientific knowledge (economic models) from the other sources of inaccuracy.

Therefore there is a strand of literature that reviewed the estimates of regulatory costs and benefits contained in IAs of environmental and occupational safety regulation in order to identify patterns of bias. Starting from regulatory costs, Hammitt (2000) compared the marginal cost of limiting chlorofluorocarbon consumption in the United States with retrospective estimates based on realised market prices. He found that estimates of compliance costs estimates were substantially overestimated, especially when compliance required the innovation and diffusion of a technology not yet available. This source of inaccuracy is often difficult to take into account. Harrington et al. (2000) provided a meta analysis of 28 cases of *ex ante* cost estimates of regulatory agencies and compared them with *ex post* cost provided by academics or independent analysts. Through a qualitative approach (*ex ante* estimates were accurate if they fall in the range of plus or minus 25% of the *ex post* assessment), they concluded that there is an overall overestimation tendency and provided useful methodological recommendations for estimating regulatory costs.

Another strand of literature focused on providing more practical recommendations on models for estimating costs and benefits. Hammitt discussed the advantages and disadvantages of different modes for valuing mortality risk (Hammitt 2000b), and contrasted adjusted life years and willingness to pay (Hammitt 2002). In a similar vein, Torriti and Löfstedt (2012) reviewed the evaluative practices related to EU IA and called for a

higher emphasis on risk analysis. They argued for the use of the value of statistical life and the price of carbon, and for an integration of macroeconomic modelling and scenario analysis. Matthews and Lave (2001) estimated occupational safety costs by relying on input-output model that allows to identify the direct and indirect economic impacts of injuries, as well as to monetise injuries, illnesses, and fatalities. Matthews (2001) proposed to assess the quality of retrospective benefit-cost analyses conducted by EPA in relation to the Clean Air Act.

Turning to methodological recommendation concerning economic regulation, to the best of my knowledge only one study evaluated the quality of IAs concerning market liberalisation. Torriti (2010) reviewed the European Commission IAs on the third package of liberalisation of energy markets. His assessment concerned also the quality of quantitative data and macroeconomic impacts. He identified several problems in the application of the chosen macroeconomic model and, more importantly, remarked the methodological flaw of utilising macroeconomic modelling in IAs that instead require the estimations to be based on individual responses to regulatory change (Torriti 2010: 1076-7). In his analysis, however, there is no reference to the best available economic knowledge on the impact of liberalisation on productivity and efficiency of energy markets. Matching the best of available scientific knowledge and methodology with the retrospective analysis of IAs is especially important with regard to EU rail liberalisation that has gone through several stages of regulatory reform. The next section will focus on the principles of economic governance in order to pursue the EU railway system (for an extensive and detailed review of EU rail liberalization see Dyrhauge 2013; Finger and Messulam 2015.

## 3 The EU rail liberalization

In December 2016, with the adoption of the market pillar of the fourth railway package, the EU single and open railway market has been finally accomplished. Railway companies operating in one member state can operate passenger services everywhere in the EU. This process of market liberalisation started in 1991 and has been gradual, typical of the European Commission's support-building strategy (Knill and Lehmkuhl 2000; Di Pietrantonio and Pelkmans 2004).

The creation of a single European railway area has been achieved through the progressive adoption and implementation of the following three principles of economic governance: i) financial separation between rail infrastructure managers and providers of rail service; ii) transparency of licensing process through the establishment of national regulatory agencies and (with the fourth package) the European railway agency that is now acting as a centralized one-shop stop for licences and safety certification of rail operators; and iii) interoperability and technical harmonisation of national rail systems.

In particular, vertical separation and the establishment of independent regulatory agencies can be considered as the most important institutional innovations associated with the EU reform (De Francesco and Castro 2016). Vertical separation requires that infrastructure managers are financially independent from railway operators. This economic governance principle was the first milestone of the EU reform. In 1991, Directive 91/440 required only the separation of account between the management of rail infrastructure and the provision of rail services. Full institutional separation has been implemented in some EU member states, whilst others have chosen a vertically integrated model where the infrastructure manager and the national incumbent rail operator are owned by a holding company. An intermediate model (in which the infrastructure manager is independent but delegates several functions to the state-owned incumbent) is also feasible. The directive also specifies that essential function such as capacity allocation, infrastructure charges

and licences must be distinct from train operators. Furthermore, the separation of account regards also rail companies operating both in freight and passengers markets.

After establishing vertical separation, the EU has pursued a gradual opening to competition and a progressive harmonization of technical and safety standards and administrative processes in order to increase the interoperability of national rail systems. Focusing on the rail freight, the first package of reforms was enacted in 2001 and defined a trans-European rail freight network. To achieve a level playing field for new entrants, this package required the independence of the national authority responsible for the licensing process from the incumbent rail operator. It is important to note that the EU railway package does not require a politically independent regulatory agency (Nash 2008: 65). In 2004, the second railway package increased the administrative transparency by furthering the specifications of freight railway interoperability and common safety standards, and by creating the European Railway Agency (ERA).

The previous regulatory reforms of freight rail paved the way for the third reform package, concerning the passenger market. The 2007 package established the service quality standard and the certification of train drivers operating within the EU, and introduced open access rights for international rail passenger services. Since 2010, cross-border rail passenger transport has been officially liberalized for all EU countries. International railway companies are allowed to pick up national traffic in a country and drop the same passengers on a further stop in the same country before crossing borders. In 2016, the fourth railway package extended open access rights to national (passengers) markets and completed the EU railway system. This package enhances impartiality between, prevents discrimination of railway operators. It also requires mandatory tender procedure for public service contracts. The next section shows how these institutional reforms and technical harmonisation have been assessed by economists and how different types of economics used for

assessing the impact of regulatory reform within railway market can be rated.

# 4 Rating the scientific knowledge on railway liberalisation

Economists have been analysing the impact of market liberalisation on productivity and efficiency gains of railways in developed countries. They rely on several indexes in order to make inferences about a firm, a set of firms, or an industrial sector (Oum, Waters, and Yu 1999). An assessment of the impact of liberalisation on railway productivity or efficiency requires isolating the differences in regulatory environments from the different sources of performance enhancement, such as organisational efficiency, economies of scale, and network characteristics, as well as exogenous factors such as technological change (Oum et al. 1999: 10). Given the methodological complexity stemming from an industrial sector with multiple outputs, economists have tended to focus on technical efficiency that is the minimisation of inputs given the level of output.[1]

Oum et al. (1999) provides an excellent review of the different methodologies for measuring productivity within railway sector. These methodologies can be "rated" according to the level of sophistication of theoretical assumptions. The following classification captures also the evolution of scientific knowledge providing a useful yardstick to assess the quality of economic analyses summarised in the European Commission's IAs and evaluation studies.

## 4.1 Indexes of productivity and efficiency

The simplest methodology for assessing productivity and technical efficiency relies on indexes that are ratio-type productivity/efficiency and does not require any statistical estimation of a production or cost function.

[1] The allocative efficiency (resulting from employing inputs in the right proportions by taking into account the rate of substitution) is impossible to measure without the knowledge of input prices (Oum et al. 1999: 13).

These indexes can be classified according to three general categories: partial and total factor productivity and data envelopment analysis method (DEA)

### 4.1.1 *Partial factor productivity*

A measure of partial factor productivity links a specific output to a single input factor, e.g., revenue tonne-kilometres per employees. Because it is easy to compute and understand, this type of productivity measures has been widely used by academics as well as the industry. Since 1981, partial factor productivity has been used to compare the performance of railway systems (Nash 1981). Comparative analyses of railway systems and firms assessed the productivity of labour, fuel and rolling stock. There are two flaws in this methodology. First, there is a problem of comparability of inputs since they are interdependent. Second, partial output index cannot represent the true total economic output whatever measure is taken (Oum et al. 1999: 14). However, the most recent partial productivity analyses combine a set of measures for distinguishing the operational and financial performances. This allows to have a gross assessment of the changes in revenues and costs (Oum et al 1999: 14).

### 4.1.2 *Total factor productivity*

A total factor productivity measure is a ratio of a total output index to a total input index. This index and can be better applied to the multi-output multi-input nature of the rail industry than partial index. The total input index is either 'the weighted sum of the growth rates of the individual input quantity indices', or the ratio of 'total expenditures (including capital) by an aggregate input price index' (Oum et al 1999: 17). The total output index is 'the weighted sum of output categories' that are freight ton-miles and passenger-miles (Oum et al 1999: 18).

There are different procedures for deriving such comprehensive indexes. In order to compare firm-level productivity instead, the methodological procedure is to assume that operating environments and economies of scale are the same across firms. This procedure has been

used since 1985 (K. D. Freeman 1985). There is a statistical method for relaxing this assumption. Decomposition regresses the total output index on 'various combinations of variables including route miles, average trip length, average length of haul, and firm dummy variable' (Oum et al 1999: 22). Decomposition has been also used by Gathon and Pestieau (1995) in the assessment of the European railways by isolating managerial efficiency from regulatory components.

### 4.1.3 *Data envelopment analysis*

Data Envelopment Analysis (DEA) is a method for constructing index of efficiency based a non-parametric linear production frontier. It requires intensive data collection of each railway company in the sample for each year of the observation time period. The DEA index varies between 0 and 1 and is dependent on the observed best practices in the sample that lie along the production frontier (Oum et al 1999: 24). Accordingly, DEA is used for benchmarking the productivity of railway firms (for instance Bookbinder and Qu [1993] ranked two Canadian and five American railway companies).

This methodology is common in network economies and has been used in railway since 1992, when Oum and Yu (1992) measured productivity efficiency in 19 OECD countries. Studies use statistical regression models to isolate the effects of different operating environments in order to quantify the effect of public subsidies and regulatory reform on firm efficiency (Oum and Yu 1994).

The main problem with DEA is related to its sensitiveness to outliers and measurement errors. Furthermore, DEA efficiency indexes are sensitive also to selection of inputs and output included in the analysis (Oum et al. 1999: 24). These limitations led economists to rely on parametric models to estimate a production function of railway operators.

## 4.2 Conventional econometric methods: deterministic or stochastic production frontier

Econometric methods estimate a production function and quantify changes in productivity or efficiency also taking into account sources and levels of inefficiency. The inefficiency term can be assumed to be a deterministic or a stochastic value. While firm dummy variables and firm-specific time trend variables are common methods for deterministic frontier models, in stochastic models the inefficiency term is the deviation of each firm from the stochastic production frontier (Oum et al 1999: 31). Both methods have been used to assess the impact of vertical separation of rail network infrastructure from operations on economies of scope in the European railway systems and companies. Stochastic methods have been to assess the impact of regulatory reform. Gathon and Perelman (1992) showed that managerial autonomy from public authority increases the technical efficiency of railway companies. Whilst based on a panel data of 12 European state-owned railways between 1973-1990, Cantos Sanchez (2001) supported the argument for vertical integration between infrastructure and operations of state-owned companies in order to avoid possible inefficiencies.

Overall, stochastic frontier method has emerged as the methodological standard for assessing the impact of the staged EU liberalisation on railway efficiency growth (Wetzel 2009). Is this methodological standard recognised in the European Commission IAs on the liberalisation of the railway? The scientific methodology for evaluating the impact of liberalisation of railway has evolved from partial and total productivity indexes to DEA-based efficiency indexes, from the use of indexes to estimations of productivity or efficiency changes. The next section assesses whether this gradual evolution has been followed also in the quality of the scientific knowledge and the methods employed in IAs on EU railway liberalisation.

# 5 Scientific knowledge and methodological sophistication in the IAs on EU railway liberalisation

## 5.1 Sample of IAs

Since 2003, the European Commission has been producing IAs on regulatory proposals. The Regulatory Scrutiny Board maintains an electronic repository of all IAs so far conducted. In such a repository, excluding I could identify 11 IAs specifically related to the EU railway regulation (see Table 1).[2] An additional IA was retrieved from the DG Move websites.[3] In this sample, there are nine IAs related to EU railway liberalisation (indicated in bold in Table 1). In the webpages of the DG Move dedicated to studies, I retrieved other seven studies associated with EU rail liberalisation packages (see Table 2).

Table 1 shows that the evaluation models utilised for assessing the economic impact of EU railway reform vary. There are six IAs that relied on qualitative assessment and multi criteria analysis; while only three IAs relied on quantitative evaluation such as dynamic model of operation ratio, regression models and financial formula. From this population of nine IAs, I have selected two IAs for in-depth content analysis. One concerns with the passengers railway reform by selecting the 2004 IA on the development of the EU railway (SEC(2004)236), the other IA, the 2010 IA on recasting the first package, concerns with the reform of freight sector (SEC(2010)1042). These IAs are related to the most important liberalisation packages concerning the freight and the passengers sector. Although limited, this sample of IAs is spread across time (2004 and 2010) and is representative of two different stages of regulatory reform. The 2004 IA was drafted in relation to the 2007 third liberation package; the 2010 IA refers to the 2012 recasting measures (intermediate liberation

[2] I have excluded from the sample IAs concerning the Trans-European Transport Network.
[3] https://ec.europa.eu/transport/modes/rail/studies/rail_en and https://ec.europa.eu/transport/facts-fundings/studies_en

steps taken between the third and the fourth package. Finally, these IAs utilise two different quantitative indexes and models.

### Table 1: IAs on EU rail regulation and liberalisation

| Year | Source | IA title | IA code | Economic impact model |
|---|---|---|---|---|
| 2004 | IA website | **Directive amending Council Directive 91/440/EEC on the development of the Community's railways** | SEC(2004)236 | Dynamic model of operation ratio |
| 2006 | IA website | **Certification and security in railway transport and interoperability of the Community rail system** | SEC(2006)1641 SEC(2006)1642 | Qualitative multi-criteria assessment of several options for each measures |
| 2007 | IA website | **Communication on rail freight oriented network** | SEC(2007)1324 | Qualitative assessment of feasibility of 4 options |
| 2007 | DG Move | Impact assessment study on rail noise abatement measures addressing the existing fleet | | |
| 2008 | IA website | **Communication on multi-annual contracts for rail infrastructure quality** | SEC(2008)132 | Multi criteria analysis with estimations of cost savings and other estimated measures of economic impact |
| 2008 | IA website | **Community guidelines on state aid for railway undertakings** | SEC(2008)517 | Quantitative financial formula for estimating the level of indebtedness |
| 2008 | IA website | **Proposal for a Regulation concerning a European rail network for competitive freight** | SEC(2008)3028 | Dynamic model using tools for transport forecasting and scenario testing[4] |
| 2010 | IA website | **Proposal for a Directive establishing a single European railway area (recast)** | SEC(2010)1042 | Regression models |
| 2013 | Register of EU COM | **Proposal for a Directive of the European Parliament and of the Council amending Directive 2012/34/EU of the European Parliament and of the Council of 21 November 2012 establishing a single European railway area, as regards the opening of the market for domestic passenger transport services by rail and the governance of the railway infrastructure** | SWD(2013)13 | Qualitative assessment of the policy options supported by quantitative elements |

[4] https://ec.europa.eu/jrc/en/scientific-tool/trants-tools-transport-forecasting-and-scenario-testing

| | | | | |
|---|---|---|---|---|
| 2013 | IA website | Proposal for a Regulation establishing the Shift2Rail Joint Undertaking | | |
| 2017 | IA website | **Proposal for a regulation on rail passengers' rights and obligations (recast)** | SWD(2017)318 | Comparison of policy option scenarios based on multi criteria analysis based on railway companies' predicted costs |

Table 2 shows that also the evaluation studies relied on a range of methods. This variation is essentially due to the different consultancies tendered for the evaluation reports. Overall, these methods both for IAs and evaluation studies are different from productivity and efficiency indexes and models utilised by economists.

Table 2: IAs on EU rail regulation and liberalisation

| Year | Evaluation title | Evaluation method |
|---|---|---|
| 2008 | Preparatory study for an impact assessment for a rail network giving priority to freight | Micro and macro – level impacts based on changes of quantitative and qualitative measures consequent to policy options |
| 2009 | Separation of account of railway undertakings and rail infrastructure managers | A study on the extent of compliance with the separation of account as required by directive 1991/440. Based on operational and financial data at railway company level |
| 2010 | Evaluation of the implementation of Regulation (EC) No 881/2004 of 29 April 2004 establishing the European Railway Agency: Results of the stakeholder analysis | A study on stakeholders' perception of the effectiveness of the ERA in rail liberalisation |
| 2010 | Study on Regulatory Options on Further Market Opening in Rail Passenger Transport | Regression analyses of the impact of market opening on increasing the rail modal split |
| 2011 | Evaluation of Regulation 881/2004 establishing the European Railway Agency (ERA) | Mainly stakeholder consultation and interviews ERA management and staff, independent analysis |
| 2012 | Further action at European level regarding market opening for domestic passenger transport by rail and ensuring non-discriminatory access to rail infrastructure and services | Qualitative assessment of regulatory options and quantitative assessment for each option of the predicted net present value (NPV = change in revenue – change in operating costs) of a standard railway company |
| 2012 | Impact assessment support study on the revision of the institutional framework of the EU railway system, with a special consideration to the role of the European Railway Agency | Quantitative measures such as NPV of the direct impacts of the different policy options |

## 5.2 An qualitative overview of the two IAs

The 2004 IA is composed of 35 pages, while the second IA is composed of 38 pages but comes with a 140 pages of appendix. The study is a massive analysis composed of a main document of 235 pages and appendixes of 305 and 405 pages. The latter appendix contains country fiches for all the 25 EU member states (Cyprus and Malta have no railway). Through a general overview of the scientific evidence contained in the two IAs, I can conclude that the above-mentioned IAs cannot stand independently since both of them rely on external analyses and studies written by consultancies. Indeed, although they estimate the impact of regulatory proposal, they do not provide sufficient information on methodology and assumptions and there is no reference to scientific and academic literature, but rather on other consultancy reports such as the one drafted by OGM titled *Developing EU International Rail Passenger Transport: Assessment of the actual and potential market for international rail passenger services*.

The 2004 RIA is based on a 165-pages study on the analysis of the impact of EU passenger rail liberalisation, drafted by Steer Davies Glaeve (2004), a leading UK-based transport consultancy. This report is freely available on the Internet. Overall, based on a EU project funded under the Framework Programme 4, this consultancy report is clear in its methodology and assumptions but provides no scientific or academic references.

The 2010 IA is also based on a consultancy report drafted by Price Waterhouse Cooper (PWC). Although cited in the IA with a webpage address, the document of the report is not available on the Internet as the provided link is not available anymore. In addition to the PWC report, the 2010 IA cites a 150-pages long study conducted by SDG on the implementation of the first rail package (Steer Davies Gleave 2005). Furthermore, this 2010 IA widely refers to data contained in another IA drafted for the European Commission Communication on railway (SEC(2008)3028). This practice of policy appraisers to rely and refer to

previous documents, consultancy reports and previous IAs creates a dispersed web of data and knowledge, rather than consolidating the organisational knowledge gained over time. Accordingly, it is impossible for a reader of these IA to understand what is the state of the art of knowledge and evidence on the impact of an ongoing process of economic liberalisation of railway in Europe.

## 5.3 Assessing the scientific knowledge within two IAs

This section is based on my careful reading of the IAs reports. My overall assessment is that while the scientific knowledge and the economic literature summaries in Section 4 are concerned with efficiency of individual railway operators or national railway systems, the two IAs on the railway liberalisation packages covered a variety of economic impacts. For instance, the 2004 SDG report underlying the first IA compares alternative policy scenarios according to the following dimension of welfare improvement: the volumes of passenger-km, the level of service provided to passengers, the fares paid by passengers, and the viability of the railway undertakings. This methodological approach to analysis a range of economic impacts beyond the rail companies is in line with the IA guidelines and the standard notion of welfare economics.

However, it is important to recall here that fostering the competitiveness of the rail systems (vis-à-vis other transport systems) and the viability of railway companies was the main priority of the EU railway reform and any initiative to liberalise utility markets. And next subsection shows that when the focus was on the economic impacts on railway operators, the methodological approaches utilised by economists are not fully utilised in IAs.

### 5.1.1 *The 2004 IA (SEC(2004)236)*

The 2004 IA and the underlying consultancy report drafted by SDG are founded on the "operating ratio", i.e., 'the ratio of the expected revenues to expected costs', of railway undertakings. Accordingly, the

model attempts to estimate the expected revenues and costs. This ratio is calculated on a set of assumptions about public service contracts, modes of tendering, and other specifications of the economic governance of railway (non-discriminatory practices in ticketing and no further regulation on rolling stock).

Forming the overall simulation (dynamic) model of the viability of the railway companies, this set of assumptions is necessary for the construction of the operating ratio. In order to assess the viability of railway operators, the consultancy report and consequently the IAs relied on a dynamic simulation model of the likely costs and revenues, profit levels and performance targets of railway undertakings. This methodological choice is sensible to the complexity to model railway undertakings' behaviour and data reliability. Specifically, the model assumes that railway companies monitor the current and the expected rates of expenditure and earnings and accordingly are able to react if the ratio of revenue falls below their profit target, by either increasing revenues or by reducing costs (Steer Davies Gleave 2004).

Overall, in order to assess the impact of liberalisation on railway companies' economic performance, economists prefer to rely on indexes of productivity or efficiency gains rather than revenue/cost ratio (cf. Section 4.1.1, revenues and costs are taken into account only through partial factor productivity). The underpinning assumptions of the former are less sophisticated but parsimonious and rely on railway companies' data that is available. Instead, revenue/cost index requires extensive data for demand-side assumptions, supply-side costs and operational decision parameters and preferences. The problem of data collection was insurmountable, as it has been acknowledged in the same IA:

> The Commission has requested several consultancy firms to assess aspects of the railway markets, but it turned out to be difficult, if not impossible, to obtain reliable figures on international passenger transport by rail, such as number of passengers; pkm; turnover; profitability, etc. Railway undertakings are reluctant to provide

> these data by invoking the commercial nature of the information. (European Commission 2004: 11)

In order to face this recognised 'difficulty of the modelling exercise' (European Commission 2004: 4), the methodological choice was to rely on '[t]he importance of the qualitative approach for the assessment, which consisted mainly of a thorough and extended survey amongst the main stakeholders, particularly the present monopolists, reveals a mixed support for a proposal for market opening' (European Commission 2004: 4). A further in-depth case study was commissioned by the European Commission to have an overview of the impact of the gradual opening up of the market for international passengers. And the selected countries were Germany, Hungary, Spain, and Sweden. This is another proof of the lack of reliance on the policy evaluation methodologies developed by economists.

### 5.1.2 *The 2010 IA (SEC(2010)1042)*

The 2010 IA refers to the recast and simplification of previous liberalisation measures contained in the first reform package, the economic analyses concerned key aspects of EU regulatory reform such as accounting separation, measures for avoiding discriminatory treatments toward new entrants and the establishment of independent (from railway infrastructure manager or undertaking) regulatory agencies for ensuring transparency in the economic governance of railway markets.

In its long appendix, this IA provides a set of impact analyses to foresee market development, the viability of railway companies and administrative costs resulting from five out of nine liberalisation proposals for facilitating market entry and competition. Four measures have been previously evaluated in a prior IA, the SEC(2008)3028, concerning a European rail network for competitive freight (see Table 1). The IA is complemented with a summary of stakeholders consultation conducted in order to define the problem, assess the effectiveness of regulatory options, and collect data for establishing the baseline scenario.

The evaluation of policy options associated with each of the five new regulatory proposals was qualitative. It combined "scores" of stakeholders and an (not specified) "independent assessment" scores through a qualitative multi-criteria analysis of implementation effectiveness. Accordingly, the selection of options was not based on welfare economics and scientific evidence. This was also attested by two resubmissions required by Impact Assessment Board that requested further improvements in the IA. For instance, the shoot out of the option of politically independent regulatory agencies is justified by the statement that this option would not increase the independence from market incumbents and accordingly the expected impacts would be exactly as the option of an agency independent from railway operators. However, scientific knowledge have already argued for the positive impact of politically independent regulatory agencies on market efficiency (Friebel, Ivaldi, and Vibes 2010; Wetzel 2009).

The preferred option was then analysed through quantitative regression models in order to quantify the economic, social and environmental impacts of the entire proposed regulatory reform. The regression analysis relied on the predicted change of several subindexes of Rail Liberalisation Index (Kirchner 2007) as the main independent variables, representing the removal of barriers to entry of new operators. Modal share of rail freight, number and market share of non-incumbents and operating cost per train/km were considered the dependent variables of regression models. The qualitative scores of the extent of effectiveness of the proposed measures are used to weight the causal direct link between barrier removal and freight rail competitiveness. Finally, a sensitivity analysis was conducted to take into account variations in the baseline scenario of the modal share of rail in freight transport.

The model identified also indirect impacts of the change of market opening consequent to the new regulatory measures. The results of the models are the followings: An increase of competition in the freight rail market attested by an increase of the new entrants and new entrants'

market share. The IA estimates an increase of recurrent administrative costs borne by public authorities (16.23 Million of Euro) and private companies (11.89 Million of Euro). The IA estimated a positive social impact attesting an increase of one thousand additional workers employed in the sector, a positive environmental impact in term of air quality, noise emission and energy consumption.

Overall, the methodological model utilised for assessing the impact of recast of the EU freight rail relies heavily on subjective judgments about the causal link between barrier removal and competition. Furthermore, there is an issue with extraneous variance since the model does not control for other possible determinants increasing the level of market competitiveness. The choice of the country level of analysis is also debatable since economist studies on the impact of regulatory changes prefer the firm level, focusing on the operative efficiency. Again there is no attempt to utilise any of indexes generally used in the economic literature. Neither is there any justification for the failure to consider the existing scientific knowledge.

## Conclusion and avenues for further research

This paper puts forward the methodological recommendation that the knowledge produced in IAs needs to be matched against the scientific knowledge and methodology. This comparison allows me to contribute to the ongoing discussion on methodological standards of *ex post* IAs that is mainly based on scorecards and economic analyses for verifying the *ex ante* estimations of regulatory costs and benefits. By applying this methodological recommendation to EU rail liberalisation, this paper present a nuanced evaluation of the practices within evidence based policy making. The empirical evidence can be summarised as follows:

1. While economists and scientific papers rely on productivity and economic efficiency and take into account the complexity of collecting data of railway companies, the knowledge produced in IAs

and evaluation studies of EU Commission rely on a vast range of evaluation methods rarely used by economists.

2. There is a remarkable mismatch between science and policy in the practices of generating knowledge. The scientific knowledge has evolved over time following a progressive pattern: from simple and partial ratios of productivity to DEA models, from panel data and deterministic model to time-series and cross-sectional applications of stochastic models of the production frontier. On the other hand, the pattern of the knowledge of IAs is scattered. Each IA (and associated consultancy report) tends to reinvent the knowledge on the impact of railway liberalisation proposing *ad hoc* economic impact methodology. This is possibly due to the plug and play effect of evaluative methods of consultancies that produced evaluation and IA studies.
3. Economists privilege parsimonious models which takes into account the availability of data; the two IAs analysed rely on either on dynamic models which requires a large set of assumptions and (unavailable) data or evaluative methods based on subjective judgements.
4. No IA summarised the state of the art and the best available knowledge of rail liberalisation. Over time, the knowledge produced throughout successive IAs is not consolidated. This undermined the legibility of IAs and the learning associated with different stage of regulatory reform.

Turning to the methodological contribution of this paper, although the research design for conducting such an assessment is straightforward, the actual development of the research has been far more complex. Specifically, the collection of evidence has been made difficult because of the absence of practice of communicating and summarising the extent and quality of scientific knowledge. Furthermore, there is a tendency of contracting out knowledge production and each consultancy puts forwards

a novel appraisal model and evaluation methodology. The richness of empirical findings confirms the soundness of the proposed evaluation methodology that assesses the extent of the difference in the practices of producing knowledge. The difference in the practices and the lack of use of the best available knowledge and evaluative methods are so large that it seems appropriate to recommend specific policymaking guidelines on how to ensure that scientific knowledge is effectively transferred in the practices of IAs.

Another methodological contribution regards the unit of analysis. In order to assess the quality of IA knowledge and the extent of reliance on the best available scientific knowledge, it is essential to analyses a set of economic analyses associated with a specific regulatory reform. IAs can be barely analysed as a standing alone document. This is because liberalisation programme come in different but connected packages, but also because IAs are often nested inside one another. This casts a doubt on previous researches that scored either a small or large sets of individual IAs.

The evident methodological limitation of the small sample of IAs paves the way for two avenues for future research. The first avenue could apply the proposed methodology to other EU liberalisation programme in order to generalise the empirical finding of this paper. The second avenue concerns the improvement of our understanding of the reasons of the barriers to the transfer and/or translation of scientific knowledge to EU policymaking and liberalisation programmes. This would require in-depth case studies that include interviews and discussions with economists and scientists in order to discern the influence of consultations, stakeholder interactions, policy-science interfaces, the role of consultancy firms within EU policy making and the extent of networked knowledge production.

## References


Adelle, Camilla, Julia Hertin, and Andrew Jordan. 2006. 'Sustainable Development "Outside" the European Union: What Role for Impact Assessment?' *European Environment* 16(2): 57–72.

Bizer, Kilian, Sebastian Lechner, and Martin Führ. 2010. 'Improving the Integrated European Impact Assessment?' In *The European Impact Assessment and the Environment*, eds. Kilian Bizer, Sebastian Lechner, and Martin Führ. Springer Berlin Heidelberg, 1–58.

Black, Julia, Martin Lodge, and Mark Thatcher. 2005. *Regulatory Innovation: A Comparative Analysis*. Edward Elgar Publishing.

Bookbinder, James H., and Wendy W. Qu. 1993. 'Comparing the Performance of Major American Railroads'. *Journal of the Transportation Research Forum* 33.

Cantos Sánchez, Pedro. 2001. 'Vertical Relationships for the European Railway Industry'. *Transport Policy* 8(2): 77–83.

Cecot, Caroline, Robert Hahn, Andrea Renda, and Lorna Schrefler. 2008. 'An Evaluation of the Quality of Impact Assessment in the European Union with Lessons for the US and the EU'. *Regulation & Governance* 2(4): 405–424.

De Francesco, Fabrizio. 2013. *Transnational Policy Innovation : The OECD and the Diffusion of Regulatory Impact Analysis*. Colchester: ECPR Press. http://strathprints.strath.ac.uk/42432/ (April 11, 2014).

De Francesco, Fabrizio, and Graziella Castro. 2016. 'Beyond Legal Transposition: Regulatory Agencies and de Facto Convergence of EU Rail Liberalization'. *Journal of European Public Policy* 0(0): 1–20.

De Francesco, Fabrizio, and Claudio M. Radaelli. 2007. 'Indicators of Regulatory Quality'. In *Regulatory Impact Assessment: Towards Better Regulation?*, eds. C. H. Kirkpatrick and David Parker. Edward Elgar Publishing.

De Francesco, Fabrizio, Claudio M. Radaelli, and Vera E. Troeger. 2012. 'Implementing Regulatory Innovations in Europe: The Case of Impact Assessment'. *Journal of European Public Policy* 19(4): 491–511.

Di Pietrantonio, Loris, and Jacques Pelkmans. 2004. 'Economics of EU Railway Reform, The'. *Journal of Network Industries* 5: 295.

Dyrhauge, Helene. 2013. *EU Railway Policy-Making: On Track?* Palgrave Macmillan.

European Commission. 2004. 'Extended Impact Assessment on the Proposal for a Directive of the European Parliament and the Council, Amending Directive 91/440/EEC on the Development of the Community's Railways to Gradually Open up the Market for International Passenger Services by Rail'.

Finger, Matthias, and Pierre Messulam. 2015. *Rail Economics, Policy and Regulation in Europe*. Edward Elgar Publishing.

Freeman, K. D. 1985. 'THE TOTAL FACTOR PRODUCTIVITY OF THE CANADIAN CLASS I RAILWAYS, 1956-1981'. *LOGISTICS AND TRANSPORTATION REVIEW* 21(3).

Freeman, Richard, and Jo Maybin. 2011. 'Documents, Practices and Policy'. *Evidence & Policy: A Journal of Research, Debate and Practice* 7(2): 155–70.

Friebel, Guido, Marc Ivaldi, and Catherine Vibes. 2010. 'Railway (De)Regulation: A European Efficiency Comparison'. *Economica* 77(305): 77–91.

Fritsch, Oliver, Claudio M. Radaelli, Lorna Schrefler, and Andrea Renda. 2013. 'Comparing the Content of Regulatory Impact Assessments in the UK and the EU'. *Public Money & Management* 33(6): 445–52.

Gathon, H. -J., and P. Pestieau. 1995. 'Decomposing Efficiency into Its Managerial and Its Regulatory Components: The Case of European Railways'. *European Journal of Operational Research* 80(3): 500–507.

Gathon, Henry-Jean, and Sergio Perelman. 1992. 'Measuring Technical Efficiency in European Railways: A Panel Data Approach'. In *International Applications of Productivity and Efficiency Analysis*, Springer, 131–147.

Goodman, Robert, and Allan Steckler. 1989. 'A Framework for Assessing Program Institutionalization'. *Knowledge, Technology & Policy* 2(1): 57–71.

Hahn, Robert W., Jason K. Burnett, Yee-Ho I. Chan, and Elizabeth A. Mader. 1999. 'Assessing Regulatory Impact Analysis: The Failure of Agencies to Comply with Executive Order 12,866'. *Harvard Journal of Law & Public Policy* 23: 859.

Hahn, Robert W., and Paul C. Tetlock. 2008. 'Has Economic Analysis Improved Regulatory Decisions?' *The Journal of Economic Perspectives*: 67–84.

Hammitt, James K. 2000a. 'Are The Costs of Proposed Environmental Regulations Overestimated? Evidence from the CFC Phaseout'. *Environmental and Resource Economics* 16(3): 281–302.

———. 2000b. 'Valuing Mortality Risk: Theory and Practice'. *Environmental Science & Technology* 34(8): 1396–1400.

———. 2002. 'QALYs Versus WTP'. *Risk Analysis* 22(5): 985–1001.

Harrington, Winston, and Richard D. Morgenstern. 2004. *Evaluating Regulatory Impact Analyses*. Resources for the Future Washington, DC. http://www.rff.org/RFF/Documents/RFF-DP-04-04.pdf (March 31, 2014).

Harrington, Winston, Richard D. Morgenstern, and Peter Nelson. 2000. 'On the Accuracy of Regulatory Cost Estimates'. *Journal of Policy Analysis and Management* 19(2): 297–322.

Head, Brian W. 2016. 'Toward More "Evidence-Informed" Policy Making?' *Public Administration Review* 76(3): 472–84.

Holvad, T. 2006. 'Railway Reforms in European Context'. In *Colloque "En Route Vers Lisbonne"(in French, Free Translation: Forum "The Route towards Lisbon") 9-10.11.*

Kirchner, Christian. 2007. 'Rail Liberalisation Index 2007. Market Opening: Rail Markets of the Member States of the European Union, Switzerland and Norway in Comparison'. *Brussel: IBM Global Business Services*.

Knill, Christoph, and Dirk Lehmkuhl. 2000. 'An Alternative Route of European Integration: The Community's Railways Policy'. *West European Politics* 23(1): 65–88.

Matthews, Deanna H., and Lester B. Lave. 2001. 'Evaluating Occupational Safety Costs and Policy in an Input–output Framework'. *Improving regulation: cases in environment, health, and safety*: 357–379.

Matthews, H. Scott. 2001. 'Analysis of the Benefits and Costs of Clean Air'. In *Improving Regulation: Cases in Environment, Health, and Safety*, eds. Paul Stelling Fischbeck and R. Scott Farrow. Resources for the Future.

Nash, C. A. 1981. 'Government Policy and Rail Transport in Western Europe'. *Transport Reviews* 1(3): 225–50.

NNR. 2006. 'The Total Cost of Regulations to Businesses in Sweden'. http://www.nnr.se/inenglish/publications/the_total_cost_of_regulations_to_businesses_in_sweden1.html (June 12, 2007).

OECD. 1999. *Guidance for Conducting Retrospective Studies on Socio-Economic Analysis*. Environment Directorate, Organisation for Economic Co-operation and Development.

———. 2003. *Regulatory Performance: Ex Post Evaluation of Regulatory Policy*. Paris: OECD.

Oum, Tae Hoon, W. G. Waters, and Chunyan Yu. 1999. 'A Survey of Productivity and Efficiency Measurement in Rail Transport'. *Journal of Transport economics and Policy*: 9–42.

Oum, Tae Hoon, and Chunyan Yu. 1994. 'Economic Efficiency of Railways and Implications for Public Policy: A Comparative Study of the OECD Countries' Railways'. *Journal of Transport Economics and Policy* 28(2): 121–38.

Radaelli, Claudio M., and Fabrizio De Francesco. 2007. *Regulatory Quality in Europe: Concepts, Measures and Policy Processes*. 1st ed. Manchester: Manchester University Press.

Radaelli, Claudio M., and Oliver Fritsch. 2011. *Measuring Regulatory Performance*. Working Paper for the OECD Regulatory Policy Committee. http://www.oecd.org/gov/regulatory-policy/2_Radaelli%20web.pdf (September 2, 2013).

Renda, Andrea. 2006. *Impact Assessment in the EU: The State of the Art and the Art of the State*. Centre for European Policy Studies.

Shapiro, Stuart. 2008. 'Evaluating the Benefits and Costs of Regulatory Reforms: What Questions Need to Be Asked?' *Evaluation and Program Planning* 31(3): 223–30.

———. 'Politics and Regulatory Policy Analysis'. http://www.cato.org/pubs/regulation/regv29n2/v29n1-7.pdf (June 25, 2007).

Shapiro, Stuart, and John F. Morrall III. 2012. 'The Triumph of Regulatory Politics: Benefit–cost Analysis and Political Salience'. *Regulation & Governance* 6(2): 189–206.

Steckler, A. et al. 1992. 'Measuring the Diffusion of Innovative Health Promotion Programs'. *American Journal of Health Promotion* 6(3): 214–225.

Steer Davies Gleave. 2004. *EU Passenger Rail Liberalisation: Extended Impact Assessment*.

———. 2005. *Railimplement–Implementation of EU Directives 2001/12/EC, 2001/13/EC and 2001/14/EC*. November.

Torriti, Jacopo. 2010. 'Impact Assessment and the Liberalization of the EU Energy Markets: Evidence-Based Policy-Making or Policy-Based Evidence-Making?' *Journal of Common Market Studies* 48(4): 1065–81.

Torriti, Jacopo, and Ragnar Löfstedt. 2012. 'The First Five Years of the EU Impact Assessment System: A Risk Economics Perspective on Gaps between Rationale and Practice'. *Journal of Risk Research* 15(2): 169–86.

Vibert, Frank. 2004. 'The EU's New System of Regulatory Impact Assessment: A Scorecard'. In European Policy Forum.

Wesselink, Anna, Hal Colebatch, and Warren Pearce. 2014. 'Evidence and Policy: Discourses, Meanings and Practices'. *Policy Sciences* 47(4): 339–44.

Wetzel, H. 2009. 'European Railway Deregulation: Essays on Efficiency and Productivity of European Railways'. Leuphana University. http://opus.uni-lueneburg.de/opus/volltexte/2009/14171/pdf/Dissertation_Heike_Wetzel.pdf?iframe=true&width=95%&height=95% (November 28, 2012).

Wilkinson, David et al. 2004. 'Sustainable Development in the European Commission's Integrated Impact Assessments for 2003, Final Report'. *Institute for European Environmental Policy, London*. http://www.basqueecodesigncenter.net/Documentos/Noticias/C22A3812-8E77-4EF3-8DD8-6D849A4FAE42/IEEP_ExIA_report.pdf (April 11, 2014).