TCB18 – PROJ 37804

Response to the Oxera Report on TCB18-GTS

Per J. AGRELL
Peter BOGETOFT

CONFIDENTIAL
2021-01-21 ver V1.0

Disclaimer

This is a specific report by the authors, professors Per AGRELL and Peter BOGETOFT for the SUMICSID GROUP.
The methods, assumptions, conclusions and recommendations in this report represent only the viewpoint of the authors that assume all responsibility for any errors of fact or logic.

Title: Response to the Oxera Report on TCB18-GTS

Final version, confidential. Project: TCB18/37804
Release date: 2021-01-21
Version. V1.0

Sumicsid Group
Rue Maurice Lietart 56
B-1150 Brussels, BELGIUM
www.sumicsid.com

Table of contents

1. Overview

1.1 Context

1.01 This note is a comment to the Oxera report "A critical assessment of TCB18 gas", called Oxera (2020) below, released 18/08/2020 on behalf GTS participating in the gas benchmarking. Oxera (2020) draws heavily on a previous analysis for electricity, Oxera (2020b) that is referenced when differing.

1.02 The format of the note is brief as most documentation is provided in the following documents, released during the project:

1) Sumicsid (2019) *Norm Grid Development, Technical* Report V1.3, 2019-02-27.
2) Sumicsid and CEER (2019a) Pan-European cost-efficiency benchmark for gas transmission system operators, Main report V1.2, 2019-07-17.
3) Sumicsid and CEER (2019b) *Project TCB18 Individual Benchmarking Reports*, GAS TSO, V1.0, 2019-07-25. (Published for GTS, among others).
4) Sumicsid and CEER (2020) Dynamic efficiency and productivity changes for gas transmission system operators, Main report V1.0, 2020-04-14.

1.03 The outline of the response restates the main arguments of Oxera (2020) in an orange shaded paragraph. In some cases, the original statements have been summarized and reformulated without intention of changing the contents and bearing of the argument.

1.04 The response provides an open discussion in a normal paragraph, concluding in a shaded grey paragraph as to our assessment of the impact of the argument on the viability of the TCB18 benchmarking results.

1.2 Outline

1.05 In Chapter 2 we recall the principles of the study and the methodological choices made in it, as well as the differences in focus in Oxera (2020) and the TCB18 project.

1.06 In Chapter 3 we respond in more detail on the main critique raised by chapter in Oxera (2020).

1.07 Appendix A contains a simulation of panel benchmarking data with data similar to TCB18, showing the effects of DEA and SFA in presence of collinear data. This appendix serves to support some of our comments in Chapter 3.

2. TCB18 Model Conception

2.1 Scope and purpose of the Oxera work

2.01 Oxera (2020) and Oxera (2020b) can be seen as a compilation of separate sensitivity analyses, applied to each element in a benchmarking study such as TCB18. The separate comments illustrate, at best, the range and frequency of impact for changes to assumptions, parameters and data in the study. However, since the development of an alternative model (or process) are not in scope of Oxera (2020), the sections cannot be compiled to a common assessment of the model quality of TCB18. E.g., in the critique of the model specification to be composed of asset-based output parameters, no comment is made to the choice of a deterministic model (DEA) and its consequences. Elsewhere, the method choice is criticized for being an unconstrained DEA, rather than a parametric method (OLS or SFA) although this would assume equally strong assumptions of the distribution of errors or inefficiency. Thus, to clarify the fundamental choices in the TCB18 study, we here revisit the data processing, the method choice and the model specification.

2.2 Data processing and validation

2.02 Data quality is primordial for benchmarking and particular attention has been given to the design of an optimal data collection and validation system.

2.03 The principles for the data collection are to ensure full understanding of the data protocol by all project participants. In TCB18 this was implemented by separate releases of the data specifications and guides in December 2017 with several rounds of reviews and two project workshops, leading to a final release in March 2018. Specific templates in Excel were developed and also revised. The project participants had ample of time and opportunity to ask questions about the data definitions, both at the interactive workshop and on open and closed areas of the project platform. Choices of principal nature, such as the activity decomposition and the scope of the benchmarking were discussed and decided jointly with the NRAs in the CEER project steering group (PSG). It can therefore be asserted that the data protocol is well known by the project participants.

2.04 The obligation to comply with any data collection procedure for a TSO is ultimately defined and enforced by the corresponding NRA. It was therefore an important principle to pass the data collection and primary data validation through the NRA, thereby inciting commitment and awareness of the TSO operations and concerns. All data exchanges, both submissions, requests for clarifications and releases of processed data, passed over the NRA to ensure full compliance.

2.05 The primary data validation was performed by the NRAs using a specific data auditing protocol, requiring the NRA to explicitly endorse the quality of the data at submission.

2.06 The role of the consultants in the data validation was to assist in the cross validation, since some TSOs did not allow other NRAs to access their data. The consultants preformed data validation of both technical and economic data in addition to the checks performed by the NRAs. The results for the data validations, frequently resulting in questions and comments, were uploaded to the project platform. Thus, each party in the project brings a specific skill to the task, improving overall quality, independency and consistency.

2.3 Choice of method

2.07 The choice of the benchmarking method (DEA) already hints at the type of model and functional form that will be privileged: the most intuitive and natural deterministic form that explains the current data and that is consistent with existing knowledge about what drives gas transmission cost.

2.4 Model specification

2.08 Cost function modelling in gas transmission is not a new science, it is well established both in the engineering and production economic literature. Whereas the reasoning in Oxera (2020) seems to suggest a wide range of different factors of unknown influence to be investigated, the structure of a good cost model can be derived from some simple principles, proven both from engineering practice, reference network analysis and transmission system benchmarking.

2.09 A gas transmission system operator is a techno-economic system with in principle four main sources of cost:

1) Transport work (direct variable cost for the transport of gas)
2) Capacity provision (fixed and variable cost for the capacity to deliver gas)
3) Grid provision (fixed and variable cost for the connection of a grid user to the main grid at a given spatial location).
4) Customer service (variable costs for the administration of grid users, safety, training and information).

2.10 As will be discussed below, the inclusion or exclusion of the categories above are consequences of the choice of model in the study.

Transport work

2.11 The transmission system provides a transport service proportional to the volume and distance of gas transported through the system. The transport work is associated with a direct cost of compressor fuel as well as costs linked to the losses of gas in the system.

2.12 A direct volume parameter as output renders the model sensitive to the cost of gas (depending on the market conditions for its purchase) and also utilization oriented. The share of cost in direct compressor fuel depends on external demand, not under the control of the TSO. This means that for a given technical efficiency (fixed capex, fixed opex except compressor fuel), the score is proportional to the demand, the higher output, the higher score. Since this is not informative for judging managerial efficiency in gas transmission, the studies Jamasb et al. (2007), E2GAS and TCB18-GAS chose not to include delivery volume as an output. Instead, the transport work dimension was expressed as the pipeline system (Jamasb et al. 2007) and its generalization, the normgrid proxy for grid size in E2GAS and TCB18-GAS. These parameters relate to the size of the network, without depending on its utilization.

Capacity provision

2.13 The gas transmission operators provide potential services through the capacity mobilized for the customers, irrespective of its utilization in any given time (see transport work). This dimension is usually covered by parameters linked to peak load, maximum capacity, compressor power, storage volume, et c.). The capacity provision cost is primarily capex linked to the assets dedicated to the physical capacity. This dimension was covered in Jamabs et al. (2007) and in E2GAS as maximum capacity (m3/year) and in TCB18-GAS as maximum compressor power (MW). The difficulty with maximum capacity is that it is related to compressor power for adjacent grids (no cost causality), entry and exit pressure and pipe dimensions. As the former is an important shortcoming and the latter is covered in TCB18-GAS through the NormGrid parameter, the physical compressor power was chosen. Note that this parameter includes all compressors irrespective of age, engine type and control equipment, which makes it different from the compressor part in NormGrid.

Grid provision

2.14 Independent of the actual volume of gas transported or the maximum flow potentially offered, the pipeline system must connect spatially distributed load and source nodes with installations that can contain and transport gas in safe and reliable manner. The infrastructure quality of the grid in itself, i.e. the interconnection of the input and output nodes is denoted as the grid provision. Whereas an increase in the compressor capacity at one of a few stations can imply significant changes to the capacity provision, the grid is normally only expanded in smaller steps, when opening new delivery areas, storage facilities and interconnections. The grid provision is present in almost all gas transmission models in the form of unweighted or weighted circuit length of the pipelines. In TCB18-GAS the NormGrid proxy contains an engineering-cost weighted circuit length of the pipelines.

Service provision

2.15 The service provision is often measured as the number of connection points or customers when dealing with distribution system operators. However, for gas TSOs the number of customers is usually very limited as the DSO (distribution system operators) handle direct delivery to low-pressure customers. In TCB18-GAS, the number of connection points, as in E2GAS and many other studies, is included as an output to cover this dimension.

Environmental factors

2.16 The class of environmental variables contains parameters that may have a non-controllable influence on operating or capital costs without being differentiated as a client output. In this class we may often find indicators of geography (topology, obstacles), climate (temperature, humidity, salinity), soil (type, slope, zoning) and density (sprawl, imposed feed-in locations). One challenge with this class of parameters is that they may be difficult to validate statistically in a small data sample. Their role of potential complicating factors will therefore have to be validated by other studies or in a process of individual claims from the TSOs. Another challenge is that in a small dataset, the explicit inclusion of many complicating factors will put pressure on the degrees of freedom in a statistical analysis. This is also the approach we have taken in this study. We have used an elaborate engineering weight system of the grid assets to reflect the investment and operating conditions. In this way, environmental factors can to a large extent be captured by the traditional output parameters.

2.17 An additional level of environmental correction was achieved by independently letting the Sumicsid power engineers derive and list the complexity factors increasing costs from a technical perspective. Note the methodological difference between doing this step prior to the specification of the model compared to a 'data mining' approach where various factors with unknown effects are used indiscriminately as regressors in a cost function. The risk with the latter approach is to find a set of factors that may fit a particular data set by second-order or spurious correlation, but without any techno-economic rationale. Another approach could be to let the environmental factors absorb any variability in cost at the average cost function stage. For example, if particularly inefficient operators have more railroad per surface area, a regression may suggest that the cost should be attributed to the railroads rather than to inefficiency. The staged method in TCB18 avoids this problem since the magnitudes of the impact is estimated a priori without inefficiency.

2.18 The variables chosen for TCB18-GAS are then two related to transport/grid provision, enhanced to directly address environmental complexity (yNormGrid_zSlope and yPipes_Landhumidity), one covering capacity provision (*yCompressors.power_tot*) and one for the service provision (yConnections_tot). We describe each shortly below.

yNormGrid_zSlope

2.19 The NormGrid provides a Totex-relevant proxy for the total pipeline system, summing all relevant assets with weights corresponding to their Capex and Opex impact. As documented in Sumicsid (2019), certain environmental conditions influence the cost of constructing and operating the pipeline system. These factors include land use type, topography (slope), vegetation type, soil humidity, subsurface features (rockiness, stones), extreme temperatures and salinity. Extensive statistical tests revealed correlations and interaction between several of the factors, e.g., vegetation and landuse type, subsurface features and topography. The most important factor for pipelines was topography (slope class), relating to costs of construction (reinforcements, site access) and to operation (maintenance access). Most other factors, with the exception of humidity, correlates with the normalized grid slope-weighted parameter. Thus, this parameter was chosen as the primary variable, explaining by itself over 95% of the variance in Totex in robust regression.

yConnections_tot

2.20 The different connection points in the transmission grid cause certain costs of operation, metering, monitoring etc. Statistically, the sum of the connections, yConnections_tot is the preferred variable, adjusted for ownership asset by asset.

yCompressors.power_tot

2.21 The transport capacity of the transmission system is measured through the sum of the installed power of the compressors units, adjusted for ownership, irrespective of type of compressor unit and type of unit. The parameter is frequently used in international comparisons and correlates both to Capex and Opex since a higher capacity requires direct costs for operation and maintenance.

yPipes_Landhumidity

2.22 Pipeline installations are especially subject to cost increases resulting from high humidity and wet soil. This results from more expensive construction site management, drainage, isolation not accounted for in the NormGrid and resources devoted to evacuation of water for repairs and preventive maintenance of segments. It was seen that dimensioning and age were not proportional to these costs. Thus, to capture this effect a technical parameter was created from the unweighted pipeline length combined with the landhumidity factors from the GIS and engineering calculations. Increasing the explanatory value, it forms a good complement to the primary parameter for grid provision, yNormGrid_zSlope.

3. Oxera critique

3.1 Transparency [Oxera, ch 2]

3.01 Oxera (2020, Chapter 2) argues that the methodological choices in the project were not transparently presented to the participants.

3.02 Modeling decision on Returns to Scale (RTS) was not presented with evidence at W2-W4, only at W5 and in the final report.

3.03 No information was given on peer companies.

3.04 Difficult to understand construction of parameters, suggesting a NormGrid calculator in Excel.

3.05 The project transparency as regards to data, or results enabling the reconstruction of data, the degrees of liberty are limited. The policy chosen by Sumicsid in the workshop and data releases was monitored and endorsed by the CEER PSG, considering the public interest.

3.06 The returns to scale assumptions will be discussed in detail from art 3.95 below. The concepts of returns to scale and the process were discussed already at W1 (Methodological Approach, page 17). Given that the last incoming data arrived very close to W3, it is evident that no results for the RTS could be presented at that workshop. At W4, already very dense with results on the environmental modelling, the linearity (CRS) of the average cost function was presented with regression results.

3.07 The final report was designed as a relatively non-technical and clear document for the final results, not a process documentation for technical readers. The project participants found ample of information on the worksmart platform, workshop presentations, methodological notes and examples, Excel calculators and individual releases of data, results and models. The interaction was also facilitated by Q&A sessions at the workshops, HelpDesk and a specific workshop entirely devoted to methodological questions posed by the project participants.

3.08 Claim 3.03 is a direct consequence of the confidentiality of data. The decision was made explicit by the PSG also for peers.

3.09 Claim 3.04 cannot be supported. Not only were there several releases of data, specifically an Excel version of a NormGrid calculator was provided for both electricity and gas: Release TCB18_ngcalc_gas_V13.xls , Version 1.3, 08/03/2019.

3.2 Data collection and validation [Oxera, ch 3]

3.10 Oxera (2020, Chapter 3) makes the argument that the TCB18 data collection and construction process does not enable a sufficiently harmonized dataset to undertake robust cost benchmarking. The section is largely extrapolating from an earlier report on the electricity data.

Data errors [Oxera, section 3.1]

3.11 Data validation by NRAs is ineffective as they only have data from their national TSO(s). Unclear whether guidelines were followed.

3.12 Data reporting for significant rehabilitations shows that (some) project participants have misinterpreted the reporting guidelines. This could bias the results towards older TSOs.

3.13 The German TSOs were not active and the validation of the data for them was unclear. The assumptions by Sumicsid on the compressor power are unstated.

3.14 German data may be incorrect due to missing ownership data (footnote 75).

3.15 In electricity, using a 10% monte carlo simulation, an error margin of 10%-18% was obtained, this may apply to gas as well. Also, SFA in electricity did not find any inefficiency, so all estimated efficiency gaps in TCB18 gas could be statistical noise.

3.16 As a development of an earlier project (E2GAS), TCB18 has reinforced the data collection and the data validation in several aspects. The project has defined data collection standards, written guides and templates, for consultation with all project participants. The data collected by the TSO passed through several rounds of NRA validation before submission to the consultant, performing a cross validation of both economic and technical data. The data validation followed a specific protocol, which was documented for all project participants, including the German TSOs.

3.17 Some assumptions had to be made when adding the German data to the sample. These assumptions were made applying cautiousness, i.e. if anything it would let the advantage to the participating TSOs. E.g., the German TSOs reported all cost in in-scope the functions Transport and Maintenance, meaning that no indirect cost was allocated to other functions. Also, they could not use the deductions for insurances, non-grid telecommunications equipment, significant rehabilitations and allowed out-of-scope deductions in TCB18. The compressor station cost in NormGrid is linear in power (MW), hence the split of the aggregate power for the German TSO to the individual stations has no impact on the results. Any (positive) distribution could be used to obtain the same result. The missing age parameter for the German TSOs is implemented as an average age for each asset.

3.18 Thus, the allegations in Oxera (2020) in this regard are speculative extrapolations from Oxera (2020b) without any concrete evidence of errors and the text contains unsupported allegations. E.g., the allegation in 3.14 above is wrong, the ownership was corrected for German TSOs.

3.19 We find no logical reason to consider the observation that only a few TSO reported significant rehabilitations as a proof of misreporting. To our best information, no project participant objected to the interpretation when presented at a Workshop.

3.20 The section on data errors contains no concrete evidence of errors, it extrapolates from another dataset and speculates on the intentions and processes in the project. Although empty, the allegations in this section are used in subsequent sections of Oxera (2020) as proven facts to support other elements.

Choice of TOTEX [Oxera, section 3.2]

3.21 This section is essentially equivalent to the one in Oxera (2020b).

3.22 TCB18 uses TOTEX, this is only acceptable if OPEX and CAPEX are equivalent at the margin and the ratio is controllable.

3.23 The choice of choosing Total expenditure (TOTEX) as the dependent variable or input in the benchmarking is the correct choice both theoretically and practically. TOTEX is used in the regulation in a number of European countries, such as Austria, Germany, Lithuania, Netherlands, Norway, Portugal, and Sweden, in benchmarking, TFP-estimations and as basis for the revenue-cap calculations, see CEER (2017). The opinion that TOTEX is a sound basis is also shared by other stakeholders, customers and operators as reported in CEER (2018). It is also the consensus of academic researchers, see e.g. the gas transmission benchmarking model in Jamasb et a. (2008). Contrary to the argumentation in Oxera (2020), benchmarking limited to e.g. OPEX would be extremely sensitive to the exact ratio (OPEX/CAPEX) that Oxera (2020) considers as partially non-controllable. By changing from leasing to direct investment, a TSO could show radical improvements in partial OPEX efficiency, but potentially without any positive impact on overall efficiency. Oxera (2020) provides an example for a TSO leasing its grid, we agree and provide it as an example of the appropriateness of the method.

3.24 The only substantive argument in this section concerns the cost normalization differences for elements that could potentially appear in either OPEX or CAPEX. Contrary to what is stated in Oxera (2020, Table 3.1), we confirm that the time period and the inflation indeed are adjusted year by year when a pooled model or analysis is made. However, the personnel cost in OPEX is normalized using the PLICI index, whereas the CAPEX is only inflation-adjusted. The reason for this is the lack of verifiable information concerning the labor element in the investments, origin and composition. To explore the sensitivity with respect to this factor, a sensitivity analysis is included in Sumicsid and CEER (2019a, art 5.26). The labor part is assumed to be between 0% and 25% of the overall investment amount. The relative difference is shown to be minimal (<1%) for the mean score and individually in the range between (-9% to +3%). Additional analyses in Agrell and Bogetoft (2020) for electricity confirm this result in TCB18, but here also including a change of index to the general LCIS index.

3.25 The critique against the choice to TOTEX lacks substance, both in theory and practice. The simulation presented lacks relevance and is inconsistent with the premises of the stated argument since it makes OPEX and CAPEX fully independent inputs, which is non-sensical.

3.26 TOTEX is the only robust input for regulatory benchmarking, since it makes the financial and operational solutions irrelevant. TCB18 has fully explored the sensitivity with respect to labor-cost corrections in CAPEX, finally not made for the general run due to lack of verifiable data.

3.27 The partial efficiencies on OPEX at a given level of CAPEX and for CAPEX at a given level of OPEX, are presented and made available to all project participants as part of Sumicsid and CEER (2019b).

Indexation of OPEX and CAPEX [Oxera, section 3.3]

3.28 This section is essentially equivalent to the one in Oxera (2020b).

3.29 Oxera (2020) considers that the price-level differences are incorrectly adjusted for. PLICI does not consider other production factors beyond civil engineering.

3.30 There is no correction for price-levels besides direct manpower cost.

3.31 TCB18 assumes open markets for all services and goods.

3.32 Objective differences may exist due to transport costs across Europe. Investments are governed by local regulation Investments over time have had different conditions

3.33 The choice of index for input-price adjustments has a methodological and an empirical side.

3.34 Methodologically, the correction for local (potentially operator-specific) input prices is the correct approach when the said prices are exogenous and well-identified. An operator required to buy land for its assets in a specific location cannot be responsible for the overall expenditure since the location is forced by the nature of the service. In the same manner, the permanent staff of a transmission system operator must be recruited and hired in accordance with national employment conditions. On the other hand, services such as invoicing, repairs, or communication could potentially be subcontracted or outsourced to service providers in the same or in neighboring states, employing part or all of the labor force under other conditions. Likewise, whereas the land and legal cost of right-of-way are intrinsically local, the value of the equipment itself and its installation are less bound to the national price-level. Frequently, transmission system operators are the only eligible buyers of certain equipment and services in their respective countries, which means that they hardly can rely upon local suppliers to provide for their needs. An erroneous correction of input prices, such as assuming that a TSO in a low-labor cost area can also acquire e.g., compressor pumps less expensively than in a high-labor cost area, will artificially skew the benchmarked OPEX negatively for the operator, irrespective of the observed cost.

3.35 Empirically, the question at hand is whether the basis for the input price correction can be well identified or even exists. Ideally, we would desire an exogenous index for the price development for all services required by a TSO and that for each country. Naturally, such index cannot be produced due to endogeneity in most countries and also the task variation across TSOs and over time. The second alternative is then to find well-defined exogenous indexes for the services for which correction is desired. Provided that such indexes exist over sufficient time and for all countries involved in the benchmarking, the operation also requires verifiable data separated over all such indexed services. E.g., an index for civil engineering involves a certain share of administrative IT-services, for which an alternative index exists, as well as construction, maintenance, auditing, et c. In reality, the choice is better guided towards a robust and well-defined basis and the closest widely available index.

3.36 In TCB18 the choice has been made to adjust for the local salary differences using the civil engineering index PLICI from EUROSTAT. The index is exogenous, available for all countries and defined for staff-intensive services without much outsourcing within the TSOs.

3.37 Oxera (2020) argues that TCB18 should adjust for all services and for investment goods. We will analyze these suggestions in turns.

Full service-price adjustments

3.38 Oxera (2020) claims that the adjustment of labor cost is insufficient and that not enough evidence is provided to validate the hypotheses behind the methodological choice. It is not clear what type of evidence Oxera would consider necessary or relevant in this case. Assuming that TSO X is shown to buy some services more expensively than TSO Y, is this evidence of varying input prices or inefficient procurement? Detailed evidence of outsourced services in other sectors, would this be representative? The argument is tautological and ignores the purpose of the benchmarking – to provide a stable platform for best-practice performance. Ad hoc adjustments to particular conditions, legacy systems and traditions would invalidate the status of the best-practice peer, as its status might heavily depend on ad hoc assumptions of past or current conditions. In the case of market changes, opening and improved procurement, the benchmarking would no longer converge to the optimal best-practice cost, but to an arbitrary state trying to explain the past. The benchmarking in itself is not the regulatory ruling, it is the NRA using the information for reviewing the performance of the TSO that would take into consideration specific suboptimal conditions that explain its occurrence.

3.39 Thus, in TCB18 and as before, we turned the question around and invited the TSOs in the operator-specific data collection in TCB18 to provide evidence of operator- or country-specific regulations or conditions that would be lasting, material and exogenous.

3.40 The correction of input prices by general price indexes is not harmless. Without observing the origin and controllability of the expenditure, it may lead to undue protection of inefficient procurement in high-cost areas and to unfair penalties for procurement in low-cost areas. The technical and economic experts in the team have observed throughout several benchmarking projects examples of services and goods procured internationally for transmission services. It has been judged more stringent in this project to refrain from assumptions regarding the nature of outsourcing (e.g. labor contents).

Adjustments of investment costs

3.41 Oxera (2020) argues for a 100% adjustment of the investment cost using PLICI. It is claimed that this corresponds to regulatory practice, citing the PR13, ORR (2013) study and one specific example for national differences in salary in OFGEM RIIO-ED1. In the case of ORR (2013), the application is actually different: the international data is transformed to GBP (the reference currency) using PPP for a five-year horizon using a TOTEX measure in nominal value, then using the UK inflation conversion. Notwithstanding some sensitivity analyses in several countries (Norway, Germany), we note no utilization of this drastic correction in any prior international benchmarking in energy, such as ECOM+ (2003, 2005), Jamasb et al. (2007), e3GRID (2009, 2012), e2GAS (2013). The claim thus stays with Oxera alone. As above, the approach is economically dubious: a major part of the investment cost in energy infrastructure is composed of materials (steel, copper) and components manufactured by a few global suppliers. Local adjustments would assume that the entire basket of overall investments would be correlated to the labor-intensive civil engineering part, usually corresponding to about 25% of the total investment. Oxera (2020) concludes from a comparison (Fig 3.2) that price-level differences can result from using a different index on the investment cost. However, Oxera provides no rationale, nor evidence as to why there could be sourcing and procurement from non-national markets for services, investment labor and goods.

3.42 In TCB18, the impact of the choice of index is illustrated in Agrell-Bogetoft(2020) and the impact of labor-cost differences in the final report (cf. 3.24 above).

3.43 The application of price-level indexes to all OPEX and/or to all labor costs in investments is introducing an unnecessary and potentially harmful assumption about the markets for goods and services in transmission provision. The index comparisons shown in Oxera (2020) neither represent well-supported cost differences due to regulatory or fiscal rules, nor do they correspond to changes in the efficiency scores (which depend on relative changes).

Allocation of indirect costs [Oxera, section 3.4]

3.44 This section is essentially equivalent to the one in Oxera (2020b) but the conclusions are somewhat different.

3.45 The sensitivity to the allocation rule is not reported (3.4.2)

3.46 Proposed to use only in-scope cost to allocate indirect costs (3.4.3)

3.47 The causation and treatment of support cost (indirect cost in TCB18) for TSOs is a recurrent question. In some past international benchmarking, all indirect costs have been included (e3GRID, 2013), in others various allocation keys have been used to create a fair comparison.

3.48 In TCB18, calculations were made as suggested in Oxera (2020, p.43) based on no indirect costs, full allocation, individual keys and common allocation keys. The results of these various options were analyzed. The impact of indirect cost was also part of the simulations in Sumicsid-CEER (2020).

3.49 The choice of allocation basis (full scope, in-scope) is a policy issue, discussed with the CEER PSG. Although Oxera (2020), on behalf of TSO not owning an LNG terminal, recommends only using in-scope costs as a basis, the opposite argument could also be made. Indirect costs for management, HR, procurement and IT-functions could serve other regulated functions as well in the interest of an overall cost-minimization. Contrary to Oxera’s request, these TSOs may rightly claim to be subject to a bias in the benchmarking since their other activities are not considered.

3.50 The allocation policy for indirect cost is primarily a choice of principle, the impact of on the efficiency results is minimal. The project steering group decided to use a partial allocation based also on non-benchmarked activities, as these were considered as relevant and economically beneficial to the core activity.

3.3 Model development [Oxera, ch 4]

3.51 Oxera (2020) generally states that the TCB18 model development appears arbitrarily restrictive and inconsistent with the scientific literature.

Cost driver analysis [Oxera, section 4.1]

3.52 Cost driver analysis was not outlined and clearly presented (4.1.3).

3.53 The econometric results (OLS) are inconsistent, since two outputs have negative signs. The current model is not validated and should not be used.

3.54 ROLS excludes outliers, these should be analyzed.

3.55 The error terms in OLS and ROLS are assumed symmetrical and normally distributed, but if inefficiency is present there will be a skew, thus the statistical inference is inconclusive.

3.56 R-squared is not an informative measure of model quality.

3.57 The application of Lasso should not be used for determining the model size.

3.58 Sumicsid has not demonstrated evidence to show absence of omitted cost drivers from the model.

3.59 The design of the final report has been addressed in art 3.07 above, the full narrative for the cost driver analysis is provided in other documents as often is the case in final reports for benchmarking projects.

3.60 The question of model specification using a regression-based average cost function versus a frontier cost model is a classical question in benchmarking. As Oxera states, the residual can be expected to be skewed if there is substantial inefficiency in the sample. The alternative would therefore be to use a parametric frontier model, such as SFA, to assist in the model specification phase. Applying a parametric frontier model, as will be shown, requires a number of non-trivial technical assumptions (distribution of the error term, structure of inefficiency term, et c.) that in themselves affect the outcome, but also a sufficiently large dataset to perform the assessment.

3.61 In Appendix A we show a constructive simulation for a benchmarking model of a size, complexity and collinearity corresponding to that of TCB18. The model has a known efficiency level and can therefore be gauged against the fit to a true outcome. The regression coefficients (OLS) for larger models are negative for one or two of the parameters, just as in TCB18-GAS. However, as shown the model gives sound estimates in DEA, rapidly converging to the true value from above (cautiousness) whereas the SFA estimation is demonstrated as of lower precision in this case. The simulation shows that for a model like TCB18, DEA provides stable and robust results even for smaller samples, whereas the parametric SFA model is unable to derive results for smaller samples and a larger error in estimation for larger data sets. There are a number of comparisons between the frontier analysis techniques DEA and SFA, such as Ferrer and Lovell (1990), Hjalmarsson et al. (1996), Bauer et al. (1993), Reinhard et al. (2000), Kousmanen et al. (2013), Andor and Hesse (2014). These analyses advocate different methods depending on the assumptions made for the underlying data references, but the consensus is that SFA performs generally well for applications with random noise and a homogenous known technology, whereas DEA is preferred for settings with low noise and heterogenous technology due to its flexibility. Our results confirm the latter point, not questioning the relevance and performance of SFA under other settings.

3.62 The scenario with negative coefficients for normal cost drivers is common in infrastructure models for production, cost and frontier functions. In the first CEER gas transmission benchmarking by Jamasb et al. (2007) the output "units" (yCompressors.power_tot) turns out significant with negative sign in Totex and Revenue models (see Table 3-1 below). The authors still do not dismiss the models, as "…[*it*] *serves two important purposes. First, it gives the benchmark R-squared as no other model (given our variables) can achieve a higher overall correlation between inputs and outputs. Secondly, as for each output variable all the coefficients are jointly significant no output variables is completely irrelevant in a statistical sense for the determination of costs.*" (Jamasb et al., 2007, p. 44). In passing we note that the authors proceed directly to the definition of their final two-parameter model (including a non-significant parameter *Capacity* from the table below) with the brief justification "*having experimented with various model specifications we selected the models shown in [Table].*"

Table 3-1 OLS results for GTSO in Jamasb et al. (2007).

Table 15: Regression results for full Cobb-Douglas models

	O&M	Totex1	Totex2	Revenue
Capacity	0.989*	0.422	0.382	0.720^
	(0.420)	(0.383)	(0.416)	(0.428)
Delivery	-0.504	0.252	0.310	-0.124
	(0.410)	(0.374)	(0.406)	(0.417)
Mains	0.336**	0.453**	0.413**	0.456**
	(0.054)	(0.050)	(0.054)	(0.055)
Horsepower	0.042	0.033	0.130*	0.191**
	(0.052)	(0.047)	(0.051)	(0.053)
Units	0.317**	-0.146**	-0.255**	-0.270**
	(0.045)	(0.041)	(0.044)	(0.045)
Load factor	0.750	0.142	0.173	0.671
	(0.678)	(0.618)	(0.672)	(0.689)
adj. R Squared	*0.88*	*0.85*	*0.83*	*0.82*

** $p<0.01$; * $p<0.05$; ^ $p<0.10$ two tailed

3.63 The effect of negative coefficients is also well recognized in larger models in other sectors. In ORR (2013) and some preceding reports on related data, counterintuitive findings for the negative sign of an obviously cost-increasing parameter (ELEC, length of electrified lines) are reported and analyzed.

3.64 The negative signs of some parameters in OLS estimates should not be interpreted as sudden proofs of reversed cost causality, they are consequences of joint estimations of collinear parameters that individually and techno-economically are validated as cost drivers. The estimation of the DEA model is not affected by these coefficients as DEA imposes a piece-wise linear cost function that freely estimates the frontier, not a linear hull for the average cost. The suggestion to discard the model has no support.

3.65 The presence of potential outliers in the data is anticipated in the data validation, model development tools and outlier filter for the calculation of the final scores. The outcomes in terms of number of outliers in each step are documented, the individual TSOs are also informed about their classification in their individual reports. Outliers in ROLS are internally analyzed for frequency, but the occurrence as extreme point in a ROLS estimate for average cost does not necessarily imply anything for the status as outlier in terms of frontier estimates.

3.66 The model fit (adjusted R^2) is not a primary output of the analysis, no predictions or estimations are made with the average cost models. The model fit is also evaluated by standard measures such as Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and Mallow's Cp to determine the informativeness of the models.

3.67 The Lasso regressions included all available parameters at the project model specification stage. They were indeed helpful to assess the variability and regularity of the sample, without resulting in techno-economically relevant models. However, the explanatory power of some proxies (road length, ruggedness) lead the project to pursue the collection and definition of individual GIS data. The Lasso regressions cannot provide absolute limits for the model size, but indications on the required number of parameters. In the case of TCB18-GAS, the minimal models in Lasso were very small, just as the models in e.g Jamasb et al. (2007). Still, the project chose a model with a higher explanatory power.

3.68 Theoretically, there can be an infinite number of omitted variables from a model. The intervention of the modeler is limited to the systematic test of whether a set of identified and a priori relevant variables outside of the model can explain efficiency. This is what is documented in the second stage part of the report.

NormGrid construction [Oxera, section 4.2]

3.69 The underlying background data for the normgrid is subject to uncertainty, not accounted for in DEA.

3.70 Derivation of NormGrid is unclear, the shares of different assets among the TSOs vary, weights may have an impact on efficiency.

3.71 Haney and Pollitt (2013) argue that aggregation in DEA contradicts the principle of benchmarking. The model should use each asset class as a separate output. Oxera argues that weight restrictions could be used if some classes get excessive dual weights.

3.72 The estimation of the normgrid component values is not absolute, but relative. This significantly reduces the uncertainty compared to norm value approaches aiming at current absolute values. Absolute estimates of cost or value would depend on the place, time and circumstances (e.g., scope of intervention) for the projects undertaken. In a large assessment with many thousands of assets as in TCB18, it would be impossible to collect data for all assets from the same place, time and setting. Summing costs from different applications would then be introducing an uncertainty in the analysis. Relative estimations, on the other hand, relate to the scaling up of a reference asset in each category, which makes the ranking very robust. Although the absolute cost for increasing a dimension in a gas pipeline vary with respect to the material cost a given year (steel), the factor for its increase remains stable across countries. The final calibration to actual cost is then a simple scaling for readability, but it is not necessary and has no impact on the DEA scores.

3.73 Assuming the relative cost values to be mean-correct, then the large number of assets involved for each operator would also tend to reduce the variability around the mean. The inclusion of aggregation or scaling weights in benchmarking to obtain output parameters (e.g., hours, tons, products, passengers) is extremely widespread and almost inevitable in any non-trivial application. To our best knowledge, no published study of DEA has included any specific analysis on this dimension.

3.74 A model specification using the different asset classes (up to six) as separate outputs in DEA is inconsistent in Oxera's own assessment for two reasons. First, an average cost estimation of the normgrid volumes by asset class would result in one or several negative coefficients due to collinearity. Deleting the concerned asset classes from the model (in accordance with the proposals from section 4.1) would lead to arbitrary bias in the benchmarking by eliminating potentially large and important asset classes. Second, filtering the operators for asset intensity and/or using output weights would introduce technical constants that are difficult to validate, basically substituting the clarity of a production economic model for a subjective weighing of different assets.

3.75 An output specification using open dual weights has no merit, aggregation is inevitable in any asset provision benchmarking.

Environmental factors [Oxera, section 4.3]

3.76 The environmental factors are not documented in the final report.

3.77 The overall approach to account for environmental factors is inappropriate since the dual weights in DEA might not be allocated to the factor for which the environmental correction is applied, e.g., Normgrid.

3.78 The environmental adjustments ignore asset location.

3.79 The correlation between selected environmental factors (scaled by NormGrid) is misleading.

3.80 Some factors, such as density, should be used as separate outputs.

3.81 The fact that dual weights are not assigned to the outputs in the same proportions is an intrinsic consequence of the weight flexibility of DEA. The dual weights are endogenously set to maximize the score for each operator, taking into account the most "competitive" dimensions. Given that many operators share similar operating conditions, very few have ideal conditions, it is not expected to see monotonous changes of score while changing the environmental conditions. This effect is not linked to the environmental correction as such, but also applies to any output dimension that is dominated for a specific operator.

3.82 The partial correlation of the environmental factors on cost is misleading since the cost includes other elements than the grid, as well as an efficiency term. The environmental complexity factors are explicitly not derived from the actual data in order to avoid this endogeneity, they are well-established and independent engineering cost estimates for the additional cost of added environmental challenges. The approach suggested in Oxera (2020) would lead to erroneous (or no) conclusions regarding the cost causality of the environmental factors.

3.83 The environmental adjustments in the international benchmarking that Oxera (2020) refers to up until E2GAS, i.e. in Jamasb et al. (2007), ORR(2013) and E3GRID (2009, 2013), were absent or reduced to binary variable ("special conditions"). In E2GAS (Agrell et al., 2016) an item-specific locational environmental correction factor was introduced in the benchmarking. Although this contributed to a higher explanatory power for the model, the variable definition and validation problems related to the approach make it unfeasible for a larger and permanent application. The approach in TCB18 is a step forward, using detailed, public and verifiable data to determine the objective service conditions in the service areas for the operators. It is correct that the asset locations are not available and used in the environmental corrections. However, the average area approach is equitable and can be seen as a proxy for the planning challenge involved: although the pipelines in a specific country may be placed in the low valleys rather than across mountains, the difficulty of finding the right path in narrow segments and passages is not costless. The approach can also be said to represent the environmental complexity of the potential rather than actual service area, an argument that may influence the dimensioning and location of stations and assets in anticipation of future connections.

3.84 The empirical result for unit cost compared to environmental adjustments is partial: the unit cost depends also on other outputs and the level of efficiency. To avoid endogeneity in the estimation in which the environmental factors would serve as coefficients covering for various other cost effects, there was no direct calibration of the TCB18-data towards the environmental complexity factors.

3.85 The use of environmental factors, such as density, as separate output factors would be a poor idea. Without scaling on some basis (e.g., assets) and weight restrictions the factor would introduce an arbitrarily high valuation of the complexity, as documented in the E3GRID (2013) model.

3.86 The environmental complexity factors are independent expert assessments, as used in any techno-economic study. They are designed not to cover other effects from a multi-output and efficiency perspective. Detailed sensitivity analysis in Agrell and Bogetoft (2020) shows a relatively low sensitivity to the choice of environmental factors or their parameters, also confirmed by the analysis in Oxera (2020).

3.4 Model assumptions [Oxera, ch 5]

Use of DEA as method [Oxera, section 5.1]

3.87 The section with general critique of DEA as a method is new in Oxera (2020).

3.88 DEA is deterministic, should (also) use SFA as in BNetzA and in E2GAS to account for stochastic errors in data.

3.89 Even absence of results in SFA should be reported.

3.90 There have already been around 20 studies published on productivity models in gas transmission, some of which cited above and in the next section. Many of those have used DEA and/or SFA as well as other econometric formulations; Cobb-Douglas, translog, COLS, etc. to investigate the properties. In E2GAS, an evolution of the models discussed in Jamasb et al. (2007) was developed and tested with both techniques, DEA for the main estimation and SFA for validation. The results for SFA confirmed the few necessary hypotheses in DEA (returns to scale), but as shown in the Appendix the SFA formulation performs poorly for small data sets with high fit to a linear cost function, just as in TCB18-GAS.

3.91 The German incentive regulation of energy networks is defined closely by the decree ARegV, including the use of the frontier methods DEA and SFA. However, it should be noted that even under ARegV, the application to smaller dataset (e.g. gas RTO) has been made without SFA for the same reasons evoked here.

3.92 E2GAS (Agrell et al. , 2016) was the seminal study for gas transmission operators, it was not clear that the modelling approach would yield stable results for a sector that had no experience with benchmarking. As in most scientific studies, the obtained models were therefore compared to previous work and the results compared to those obtained by secondary methods such as SFA and Unit-Cost analysis. It was clearly highlighted in the analysis (Agrell et al. , 2016, art 5.35) that the use of these methods, drawing on different assumptions, should not be seen as alternatives, but merely as validation of the approach. It is customary to run more extensive method cross-validation for seminal studies, but not necessarily for repeated applications of an already established and proven method on a data from the same sector.

3.93 The differences between DEA and SFA can be illustrated with a simple example. The use of DEA as cost estimator in regulatory economics is comparable to its use in evaluation of multi-dimensional auctions in environmental protection. In this application in Denmark, farmers or their associations offer projects as 'bundles' involving acres of land using organic agriculture, protection of land and species, ground water offtake and wild forest growth in open volumes at a given price (cost). The regulator (proxy-buyer) chooses a selection of the efficient offers in DEA to allocate the subsidies in the most efficient way across the required land and volumes. Note that this is done *ex ante* as an auction, the agents are paid to achieve their offer. If SFA were to be applied to this scenario, the offers would be smoothed out since the best are considered “too good to be true” and the differences between the offers would be considered as “chance” not to be included in the auction. In consequence, almost no offer would fully win the auction, but the more different they are, the more of them would be partially selected. In an auction this makes little sense. However, using validated *ex post* data for performed services and installations put into service in regulatory economics is also assumed not to be random (as the tariffs are deterministic and the services billed for should be as well). Hence, an SFA validation might give some indication of the rank order of operators, but not of the actual magnitude of efficiency.

3.94 DEA has been chosen by CEER for the TCB18 benchmarking on cross-sectional data for its absence of a priori assumptions on the production function and the structure of the potential inefficiency in the sector. Crossvalidation with other methods is in general a good idea, provided they can produce valid and comparable results. The model specification for TCB18-GAS is based on a model structure and data that are already validated with a number of different techniques, including SFA in the two recent studies of European data. In TCB18-GAS, the model development aimed at an improved explanatory power for the new environmental factors. This was successful and useful for the DEA model intended for application. However, the higher fit reduces the noise necessary for parametric models such as SFA to determine their estimates and separating random noise from inefficiency. Validation with SFA is also not applicable in TCB18 as the model requires larger datasets to yield useful results and even so does not converge to the correct efficiency for the type of cost function use (see Appendix). Consequently, SFA validation was not judged to be informative for the models. Validation with a simple ratio method such as Unit Cost (cf. ECOM, E3GRID) is of course possible but the results would be misleading as this ignores the multiple outputs and the environmental aspects..

Returns to scale [Oxera, section 5.2]

3.95 This section is essentially equivalent to the one in Oxera (2020b).

3.96 Banker F-test and sum of coefficients in log-linear regression are not presented in the report.

3.97 Banker test: Oxera finds variable returns to scale (VRS), not non-decreasing returns to scale (NDRS). Sum of coefficients is less than 1, but not significantly (constant returns to scale; CRS).

3.98 Testing intercept in levels: weak support for decreasing returns to scale (DRS).

3.99 Sumicsid should explain why large TSOs are systematically estimated to be inefficient under NDRS.

3.100 Oxera hypothesizes that the allegedly lower efficiency of larger TSOs could be caused by non-modelled factors such as old asset age, density or general size effects.

3.101 Oxera (2020) devotes a large discussion to a question that is settled since quite a while in the academic world. Already in 1949, the young doctoral student Hollis Chenery, later becoming the World Bank’s VP for Development Policy, published a fundamental paper combining microeconomic theory with production economics and engineering. The application in Chenery (1949) was indeed the natural gas transmission industry. Based on a detailed cost analysis of just two assets: the pipeline and the compressor station, he showed how the output could be flexibly adjusted by installing more compressors for the same pipeline, by extending the network, or by the two actions. Mathematically, the significant economies of scale in gas transmission were established already in 1949. Drawing on the theoretical and analytical work, successors found empirical evidence for scale economies, such as Robinson (1972) at 2.07 and Callen (1978) at 1.17. A widely cited work by Aivazian et al. (1987) investigate the issue whether economies of scale or technological change drives productivity growth in the gas transmission industry. Based on a timeseries from 1953 to 1979, the authors show persistent and significant economies of scale across time, dominating the technological change as productivity driver. Sickles and Streitwieser (1992, 1998) and Granderson and Linvill (1996) confirm the conclusion under regulatory constraints, showing that gas transmission operators rely basically entirely on scale economies for efficiency and productivity, the technological change being driven by regulatory pricing.

3.102 Gordon et al. (2003) addresses the issue of economies of scale and scope from a fundamental viewpoint using the concept of sub additive cost functions, a condition for the quality of a natural monopoly for the industry. In short, a sub additive cost function shows decreasing marginal cost per output when volume or some size proxy is increased. Using Canadian data for gas transmission, Gordon et al. (2003) confirms the economies of scale as the dominant cause of subadditivity and the resulting natural monopoly status. Yepez (2008) is returning to the engineering roots of Chenery (1949) and looks at the cost function for interstate gas transmission, focusing on pipeline and compressor station costs, decomposed in capital, operating and maintenance costs. The paper shows consistent increasing returns to scale in all input dimensions, as well as in total expenditure. Massol (2011) confirms the engineering cost function by Yepez (2008) in showing that even a single-factor model suffices to validate the economies of scale in the sector.

3.103 Jamasb et al. (2008) discuss the links between European and US gas transmission operators, empirically validating economies of scale using two cost translog cost functions with values around 0.7 – 0.9 (meaning that a 100% increase in output would only increase cost by 70% - 90%). The study, also using DEA for a Malmquist analysis can be said to be a direct predecessor for the European regulatory study Jamasb et al. (2007) on behalf of ERGEG (now: CEER).

3.104 The abundant evidence on the economies of scale above was reconfirmed in the E2GAS project, Agrell et al. (2016) for a subset of the European gas TSOs in the TCB18 project, using an log-model with the outputs NormGrid, connection points and maximum capacity.

3.105 The returns to scale assumption in TCB18 is based on econometric tests, observations from the empirical distribution of efficiency and techno-economic considerations from the model specification. Given the number of observations in a given project and gradual model development in the fit to controllable cost, it is evident that the existing scientific and techno-economic knowledge is utilized in the modelling.

3.106 Although one may argue that the statistical documentation on returns to scale is not complete in the TCB18 report, the objection behind reveals a conjecture that is flawed. As tests for returns to scale are hypothesis-based, the approach advocated by Oxera (2020) would logically assume that even if economies of scale are proven scientifically in numerous studies and in a previous study for the same operators, a repetition with a smaller number of operators would then lead to the rejection of the hypothesis. We consider this idea contrary to a scientific approach to performance assessment.

3.107 The default approach in DEA should be to impose only the minimally necessary assumptions, to avoid a priori influence on the data and the efficiency assessment. In this case, it would correspond to variable returns to scale (VRS), meaning that there could be both increasing and decreasing returns to scale in the data set.

3.108 Besides the abundant evidence from other studies for increasing returns to scale (or to be more exact NDRS), there are also specific regulatory arguments for this choice. Experience and data also from other energy network benchmarking highlight the observation that the concessions for national transmission system operators cannot readily be extended outside of their borders, for technical, regulatory and economic reasons. It may be infeasible to gauge a small operator against a larger with the argument that full scale-efficiency should be obtained.

3.109 First, a simple look at the distribution of efficiency in TCB18 gas provides concrete evidence for the assumption. A plot[1] shows observations of high or full efficiency across all sizes, from the smallest to the largest operators. Even under NDRS (the score is almost identical for CRS, we can observe inefficient and efficient observations across different sizes. It is true that there are only four large units in the sample, but also among these units there is considerable differences in efficiency.

3.110 VRS as assumption is rejected by all published studies and the previous regulatory benchmarking, but also simple techno-economic evidence: no valid reason can be found to suggest that larger operators would be unable to organize their services and assets as efficiently as smaller units. Oxera (2020) provide no other argument for why this assumption would be valid beyond purely small sample size effects. The purported speculations regarding asset age and density have not been validated with statistical tests. Age across the sample does not significantly increase the total cost per output. Density is covered in the model through the landuse factors that were extensively tested without revealing any specific effects related to size. Finally, the idea of some generic diseconomies of scale contradicts both theory and all empirical studies in the area.

3.111 The assumption for NDRS is based on techno-economic arguments linked to the abundant evidence from other studies, previous transmission studies, concession areas and the model specification (NormGrid covers all relevant assets). In particular the conjecture that there suddenly would be diseconomies of scale in this dataset for gas transmission is not only statistically unfounded but also absurd from a techno-economic viewpoint.

Outlier analysis [Oxera, section 5.3]

3.112 This section is essentially equivalent to the one in Oxera (2020b).

3.113 A technical section largely based on Oxera work for the industry against BNetzA and the incentive regulation in Germany.

3.114 TCB18 uses outlier detection as prescribed in German law (ARegV), but Oxera has supported industry appeals against ARegV and BNetzA in that respect.

3.115 Scores are not half-normal, but DEA is non-parametric. The Banker test is of value only for large samples.

3.116 Efficiencies of the same unit in both numerator and denominator is not consistent with Banker independent samples.

[1] The plot contains confidential information concerning all TSOs and cannot be reproduced here.

3.117 Sumicsid has agreed on the objections against the F-test in other legal processes.

3.118 Super efficiency should be iterative (Thanassoulis, 1999).

3.119 Oxera (2020) is partially repeating arguments from an appeal that they were involved in on behalf of operators against BNetzA concerning the implementation of the outlier detection in accordance with ARegV. Oxera argues against dominance, super-efficiency tests and in favor of bootstrapping and/or sequential application of outlier filtering.

3.120 We will not repeat the lengthy arguments in the appeal against BNetzA here (cf. Agrell and Bogetoft, 2019), it suffices to clarify that contrary to the allusions in Oxera (2020), the appeal was rejected and the outlier detection procedure in Germany practiced by BNetzA for both electricity and gas networks at all levels remain the most advanced and best practice in regulatory benchmarking. To our knowledge and not contradicted by Oxera, no NRA uses the suggested bootstrapping procedure.

3.121 Outlier detection is a vast area of academic discussion, see also Agrell and Niknazar (2014) with various models advanced for detecting outliers, defined in various ways. As noted in CEER-Sumicsid (2019), outlier detection for the type of sample used in TCB18 is not merely a mechanical application of the criteria in ARegV, it also includes econometric reviews such as Cook's distance and foremost studies of how individual units appear in different graphs for unit costs and certain partial measures. This holistic approach, combined with the data cross validation, warrants for the greatest possible protection of the replicability of the efficient frontier. The exclusion of one TSO is based on a combination of econometric and techno-economic observations.

3.122 The outlier detection in TCB18 follows best practice for regulatory benchmarking, well beyond studies cited in Oxera (2020) such as the ORP (2013) or others in which mainly ad hoc inspection is used. Outlier detection in a small data set is always a careful multi-tool balance, not a mechanical application.

Second-stage analysis [Oxera, section 5.4]

3.123 This section is essentially equivalent to the one in Oxera (2020b).

3.124 Second-stage analysis is not correct, the second-stage parameters are not independent from the first-stage parameters, no correction for serial correlation.

3.125 The procedure does not guarantee absence of omitted cost drivers in gas,

3.126 The critique against the post-run (second stage) analysis is not well posed since the purpose is not related to model-specific variable selection. The post-run analysis aims at investigating and validating the potential presence of systematic bias for operators with specific conditions with respect to the efficiency score used in the regulation. The list of parameters includes also elements already in the model, for information about potential impact and not as an omitted variable.

3.127 The model specification process includes a structured approach for covering the services of a TSO; grid provision, capacity provision, customer service. An additional stage included systematic incorporation of environmental conditions. The post-run second-stage process is intended to detect potential bias in the scores, not the inclusion of specific parameters.

Frontier shift [Oxera, section 5.5]

3.128 Regress in dynamics means that the model is wrong, it misses changes in regulatory burden over time.

3.129 The realistic explanation for the regress is model misspecification, model does not take into account differences in efficient expenditure over time.

3.130 The dynamic report is unclear on outlier analysis and the return to scale chosen.

3.131 The caveat in the dynamic report hints at the role of sample size.

3.132 Inconsistent to assess one result (static) as more robust than another (dynamic) if both are estimated using the same model and similar datasets. Both results should be rejected.

3.133 The critique by Oxera (2020) is essentially based on two (unstated) assumptions:

1) Productivity development in (gas) transmission must be positive, unless the model is misspecified.
2) The predictive quality of a Malmquist model must be the same as that for a cross-sectional run with the same model.

3.134 Concerning the claim (1), it is fundamentally wrong from both theoretical, economical and empirical reasons.

3.135 Theoretically, it is inconceivable that a given sector would show positive productivity growth irrespective of the sample of operators, regulatory regime and length of time period. In particular in an infrastructure sector like transmission with lumpy investments, it is unlikely that new technology and processes can penetrate the operations continuously over time. A more plausible hypothesis is that new technology is implemented only in discrete steps and that the intensity of restructuring of the processes depend on the regulatory and managerial context at the time. Investments in pipelines and compressors are discrete and the efficiency is increasing with the utilization of the assets, as well as the learning and procurement effects of input scale. It is therefore logical that a major component (such as in Callen (1978)) of productivity increases are caused by capacity investments.

3.136 Empirically, our hypothesis is supported by the fundamental work on the sources of productivity growth, such as Aivazian et al. (1987) showing the general dependency on scale expansion to attain productivity growth for the early days of automation 1953-1979. During this period the pipeline network was augmented by larger diameter line-pipe and the compressor stations were converted to automatic or semi-automatic operations as well as more fuel-efficient engines, leading to fuel and labor cost savings and capacity increases. Consequently, this period was characterized by productivity increases linked primarily to technical change (contribution 53%) and secondary to returns to scale (contribution 34%).

3.137 However, Sickles and Streitwieser (1992, 1998) analyze the US interstate gas TSO for the period 1977-1985 following the introduction of new regulation, the Natural Gas Policy Act (NGPA) in 1978. Using both DEA and SFA, they find consistent proofs of declining efficiency during the period, on average the TFP regress was -1.18% per year. Granderson and Linvill (1996) use a translog cost function to analyze the same data (1977-1987), confirming the negative productivity development, but refining the analysis to consider the impact of regulatory cost caps. Since the regulation in itself initially limit the cost increases, the efficiency effect is seemingly positive. Correcting for the effects of lagged regulation caps, Granderson and Linvill (1996) find that scale economies account for 53% of the productivity growth.

3.138 In general transmission, temporary productivity regress is not rare, e.g. Llorca et al. (2016) find regress in USA for electricity transmission during the period 2001-2009 using an SFA application correcting for environmental effects. Using a TFP approach, AER (2019) reports continuous regression in electricity transmission from about 2009 to 2018, see also Figure 1. The reasons for regress are probably multiple and beyond this note.

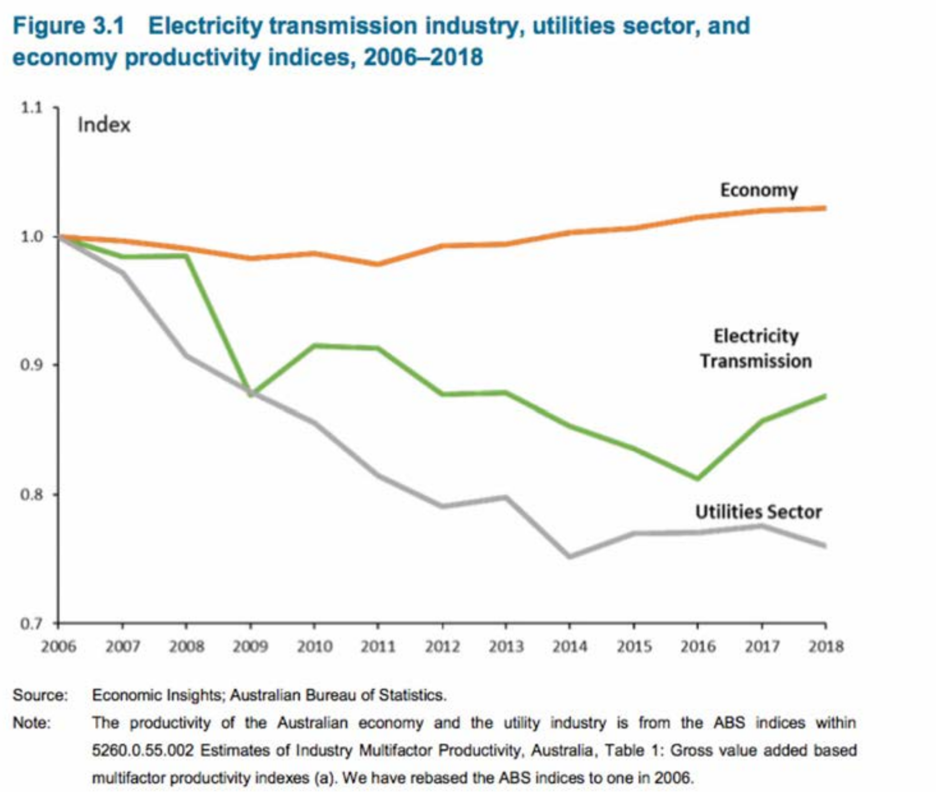

Figure 1 Total factor productivity for industry, electricity transmission and utilities, 2006-2018. AER (2019).

3.139 Topp and Kulys (2012) look at general multifactor productivity in the Australian utilities (electricity, gas, water and waste services) 1985-2010. As seen in Figure 2, the trend is negative from 1997. The analysis here points at cyclical investment patterns (see above), unmeasured output and quality changes or incremental higher cost for expansion.

Figure 2 **Electricity, gas, water and waste services: Inputs, output and multifactor productivity, 1985-86 to 2009-10**

Index 2008-09 = 100

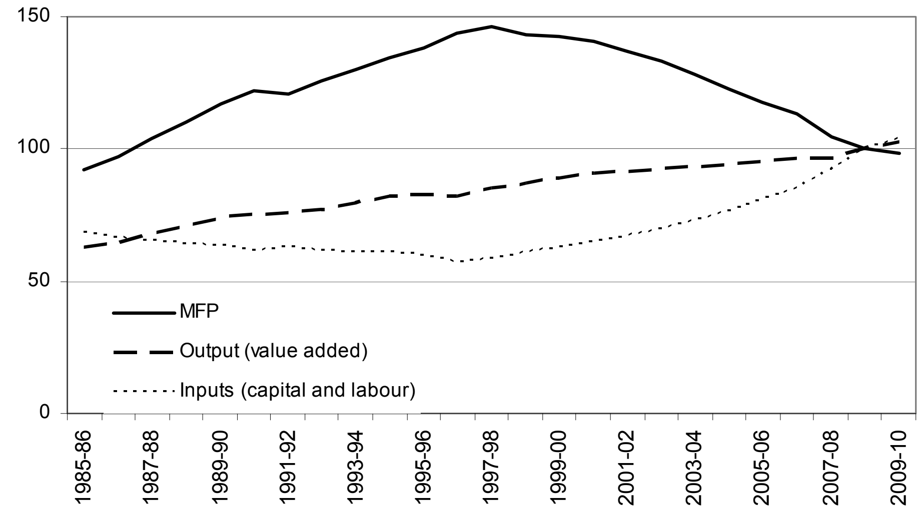

Figure 2 Productivity development utilities, Australia, 1985-2010 (Topp and Kulys, 2012).

3.140 Oxera (2020c) investigates the productivity development for the Belgian electricity and gas distributor Fluvius 2015-19 using a DEA Malmquist model with the outputs pipelines, connection points and energy delivered and constant returns to scale (CRS). The findings are presented in Table 3-2 below. As seen, Oxera (2020c) concludes on a negative frontier shift (regress) for the period, without any suggestions regarding the model validity or the data situation in the report. Note that the policy recommendation for the commissioning regulator (VREG) in Oxera (2020c) is to use a positive frontier shift derived with another model amounting to 0.4% per year.

Table 3-2 Frontier shift results in GDSO 2015-19, Oxera (2020c).

Table 4.2 Frontier shift—DEA

	2015–18	2015–19
Electricity distribution (% p.a.)	1.4%	0.8%
Gas distribution (% p.a.)	-0.1%	-2.4%

3.141 The allegation that productivity regress should be an indication of model errors is rejected by theory, economics and empirical results from a number of authors (including Oxera). Productivity changes in gas transmission have occurred in both directions and the reasons evoked by Oxera (2020) are not consistent with scientific work in the area.

3.142 The second claim relates to the differences in interpretation between the dynamic and static models and their results.

3.143 The dynamic model is a conventional Malmquist formulation under CRS, excluding outliers and TSOs for which only a single year is available. As seen in Table 3-3 below, a reproduction of Table 3-4 in Sumicsid-CEER (2020), the annual datasets estimating the frontiers have between 9 and 11 TSOs included. The findings for which a caveat is formulated concern specifically the frontier shift, i.e. the relative changes of the four-dimensional output frontier in between years. Dependent only on a few peer units, partially shifting over time, the magnitude of the changes cannot be determined with precision.

3.144 The static frontier is established by 23 TSO observations with complete data excluding 6 outliers. The data and the robustness of the results in DEA have been confirmed using several nonparametric methods and the documented sensitivity analyses.

Table 3-3 Malmquist results in Sumicsid-CEER (2020, Table 3-4).

	Malmquist	Efficiency change	Technical change	Number of DMUs
2013 - 2014	0.998	1.010	0.988	9
2014 - 2015	1.028	1.030	0.999	10
2015 - 2016	0.976	0.996	0.981	11
2016 - 2017	0.996	1.033	0.966	11
Mean	1.000	1.017	0.983	

3.145 The dynamic and static models are identical, but the number of units included in the runs are widely different (9-11 vs 23) excluding outliers, as well as the finding at hand (frontier shift vs individual operator score). Thus, the caveat for the dynamic results is as justified as the robustness of the static scores.

3.5 Oxera summary [Oxera, ch 6]

3.146 The final summary in Oxera (2020, Chapter 6) is not adding any new information.

3.147 Oxera is concerned about the statement in the final report on future work directed towards refinements rather than model development.

3.148 The recommendations in Oxera (2020, Appendix A2) contain a number of elements that Sumicsid can agree on as important for a regulatory benchmarking, although Sumicsid and Oxera may disagree on the implementation of some of these dimensions.

References

Agrell, P. J. and P. Bogetoft (2007) Development of benchmarking models for German electricity and gas distribution, Final report 2007-01-01, Bundesnetzagentur (BNetzA), SUMICSID. DOI: 10.13140/RG.2.2.21296.17927

https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Energie/Unternehmen_Institutionen/Netzentgelte/Gas/Effizienzvergleich%20Verteilernetzbetreiber/GA_Sumicsid_Effizienzvergleich.pdf?__blob=publicationFile&v=3

Agrell, P. J. and P. Bogetoft (2017) Reponse to Prof. Kumbhakar and Prof. Thanassoulis – Validity of Bundesnetazagentur's cost driver analysis andsecond-stage analysis in its efficiency benchmarking approach. Report to the Court, 2017-05-15. SUMICSID.

Agrell, P. J. and P. Bogetoft (2019) *Outliers in DEA-based regulatory benchmarking*. Response to Oxera, V1.3, 2019-02-03. SUMICSID.

Agrell, P. J. and P. Bogetoft (2020) *Sensitivity Analysis for TCB18 Results for Statnett*. Final report for NVE, 2020-05-04. SUMICSID. Confidential.

Agrell, P. J., Bogetoft, P. and U. Trinkner (2016) *Benchmarking European Gas Transmission System Operators*. Project E2GAS, Final report for CEER. SUMICSID and Swiss Economics. DOI: 10.13140/RG.2.2.31362.50880

Agrell, P. J., and Brea-Solís, H. (2017). Capturing heterogeneity in electricity distribution operations: A critical review of latent class modelling. *Energy Policy*, 104, 361-372.

Agrell, P. J., and Niknazar, P. (2014). Structural and behavioral robustness in applied best-practice regulation. *Socio-Economic Planning Sciences*, 48(1), 89-103.

Aigner, D., Lovell, C. K., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1), 21-37.

Aivazian, V. A., Callen, J. L., Chan, M. L., & Mountain, D. C. (1987). Economies of scale versus technological change in the natural gas transmission industry. *The Review of Economics and Statistics*, 556-561.

Andor, M., & Hesse, F. (2014). The StoNED age: the departure into a new era of efficiency analysis? A monte carlo comparison of StoNED and the "oldies"(SFA and DEA). *Journal of productivity analysis*, 41(1), 85-109.

Banker, R.D. (1996), Hypothesis Test Using Data Envelopment Analysis, *Journal of Productivity Analysis*, 7, pp. 139-159.

Battese, G.E. and T. Coelli (1992), Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis*, 3, 153-169.

Battese, G., T. Coelli, and T. Colby (1989) Estimation of Frontier Production Functions and the Efficiencies of Indian Farms Using Panel Data from ICRISTAT's Village Level Studies, *Journal of Quantitative Economics*, 5, 327–348.

Bauer, P. W., Berger, A. N., Ferrier, G. D., & Humphrey, D. B. (1998). Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods. *Journal of Economics and Business*, 50(2), 85-114.

Bogetoft, P. (1997), DEA-Based Yardstick Competition: The Optimality of Best Practice Regulation, *Annals of Operations Research*, 73, 277-298.

Callen, J. L. (1978). Production, efficiency, and welfare in the natural gas transmission industry. *The American Economic Review*, 68(3), 311-323.

Capeluck, E., & Capeluck. (2016). *A Comparison of Australian and Canadian Productivity Performance: Lessons for Canada* (No. 2016-07). Centre for the Study of Living Standards. http://www.csls.ca/reports/csls2016-07.pdf

CEER (2017) CEER *Report on Investment Conditions in European Countries*. CEER Conclusions Paper C16-IRB-29-03.

CEER (2018) *Incentives Schemes for Regulating Distribution System Operators, including for innovation*. CEER Conclusions Paper C17-DS-37-05.

Chenery, H. B. (1949) Engineering production functions. *The Quarterly Journal of Economics*, 63(4), 507–531.

Chenery, H. B. (1952) Overcapacity and the Acceleration, *Econometrica*, 20(1), 1–28.

Farsi, M., Fetz, A., & Filippini, M. (2008). Economies of scale and scope in multi-utilities. *The Energy Journal*, 29(4).

Ferrier, G. D., & Lovell, C. K. (1990). Measuring cost efficiency in banking: Econometric and linear programming evidence. *Journal of econometrics*, 46(1-2), 229-245.

Gordon, D. V., Gunsch, K., & Pawluk, C. V. (2003). A natural monopoly in natural gas transmission. *Energy economics,* 25(5), 473-485.

Granderson, G., & Linvill, C. (1996). The impact of regulation on productivity growth: an application to the transmission sector of the interstate natural gas industry. *Journal of Regulatory Economics,* 10(3), 291-306.

Haney, A. B., & Pollitt, M. G. (2013). International benchmarking of electricity transmission by regulators: A contrast between theory and practice?. *Energy Policy*, 62, 267-281.

Hjalmarsson, L., Kumbhakar, S. C., & Heshmati, A. (1996). DEA, DFA and SFA: a comparison. *Journal of Productivity Analysis*, 7(2-3), 303-327.

Jamasb, T., Newbery, D., Pollitt, M., & Triebs, T. (2007) *International benchmarking and regulation of European gas transmission utilities*. Report prepared for the Council of European Energy Regulators (CEER). http://portal.e-control.at/portal/page/portal/EER_HOME/EER_PUBLICATIONS/CEER_PAPERS/Gas/2006/ERGEG_Cost_benchmark_final_tcm7-107140.pdf

Jamasb, T., Pollitt, M., & Triebs, T. (2008) Productivity and efficiency of US gas transmission companies: A European regulatory perspective. *Energy Policy*, 36(9), 3398-3412.

Kuosmanen, T., Saastamoinen, A., & Sipiläinen, T. (2013). What is the best practice for benchmark regulation of electricity distribution? Comparison of DEA, SFA and StoNED methods. *Energy Policy*, 61, 740-750.

Llorca, M., Orea, L., & Pollitt, M. G. (2016). Efficiency and environmental factors in the US electricity transmission industry. *Energy Economics*, 55, 234-246.

Massol, O. (2011). A cost function for the natural gas transmission industry: further considerations. The Engineering Economist, 56(2), 95-122.

Mehdiloo, M., & Podinovski, V. V. (2019). Selective strong and weak disposability in efficiency analysis. *European Journal of Operational Research*, *276*(3), 1154-1169.

Nerlove, M. (1961). Returns to scale in electricity supply. In *Measurement in Economics – Studies in Mathematical Economics and Econometrics in Memory of Yeshuda Grunfeld*, C.F. Christ (Ed.), Stanford University Press.

Oliver, M. E. (2015). Economies of scale and scope in expansion of the US natural gas pipeline network. Energy Economics, 52, 265-276.

ORR (2013) *PR13 Efficiency Benchmarking of Network Rail using LICB*. Office of Rail Regulation.https://www.orr.gov.uk/sites/default/files/om/pr13-efficiency-benchmarking-of-nr.pdf

Oxera (2020) *A critical assessment of TCB18 gas*. Report for GTS.

Oxera (2020b) *A critical assessment of TCB18 electricity*. 2020-04-30. https://www.oxera.com/wp-content/uploads/2020/06/A-critical-assessment-of-TCB18-electricity_Oxera_Final.pdf

Oxera (2020c) *The necessity and magnitude of frontier shift for the Flemish electricity and gas distribution operators over 2021–24*, Report for VREG, Belgium. 2020-02-13. https://www.vreg.be/sites/default/files/Tariefmethodologie/2021-2024/oxera_report.pdf

Reinhard, S., Lovell, C. K., & Thijssen, G. J. (2000). Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. *European Journal of Operational Research*, 121(2), 287-303.

Robinson, S. T. (1972). Powering of natural gas pipelines. *Journal for Engineering for Power*, 181-186.

Sickles, R. C., & Streitwieser, M. L. (1992). Technical inefficiency and productive decline in the US interstate natural gas pipeline industry under the Natural Gas Policy Act. In *International Applications of Productivity and Efficiency Analysis* (pp. 115-129). Springer, Dordrecht.

Sickles, R. C., & Streitwieser, M. L. (1998). An analysis of technology, productivity, and regulatory distortion in the interstate natural gas transmission industry: 1977–1985. *Journal of Applied Econometrics*, 13(4), 377-395.

Sumicsid (2019) *Norm Grid Development*, Technical Report V1.3, 2019-02-27.

Sumicsid and CEER (2019a) *Pan-European cost-efficiency benchmark for gas transmission system operators*, Main report V1.2, 2019-07-17. https://www.ceer.eu/documents/104400/-/-/559c7df0-9cf3-2153-07bd-855bdf9a6a13

Sumicsid and CEER (2019b) *Project TCB18 Individual Benchmarking Reports*, GAS TSO, V1.0, 2019-07-25. (Published for Energinet.dk and GTS). https://forsyningstilsynet.dk/media/6260/individuel-gas-energinet.pdf and https://www.acm.nl/nl/publicaties/ontwerpmethodebesluit-gts-2022-2026

Sumicsid and CEER (2020) *Dynamic efficiency and productivity changes for gas transmission system operators*, Main report V1.0, 2020-04-14. https://www.acm.nl/sites/default/files/documents/sumicsid-ceer-dynamic-efficiency-and-productivity-changes-for-gas-transmission-system-operators.pdf

Topp, V. and T. Kulys (2012) *Productivity in Electricity, Gas and Water: Measurement and Interpretation*, Staff Working Paper, Productivity Commission, Australian Government. https://www.pc.gov.au/research/supporting/electricity-gas-water/electricity-gas-water.pdf

Yatchew, A. (2000). Scale economies in electricity distribution: A semiparametric analysis. *Journal of Applied Econometrics*, 15(2), 187-210.

Yépez, R. A. (2008). A cost function for the natural gas transmission industry. The Engineering Economist, 53(1), 68-83.

Appendix A

4. DEA and SFA with collinear data

In Oxera (2020) a number of claims are made against the validity of a model when some regression coefficients are negative, stating that this would lead to potential errors in a DEA model, as well as a series of statements about the use of SFA instead of DEA to obtain robust results.

The argumentation is fragmented although each point can be met theoretically and technically, it may leave the reader with the impression that the model is somehow deficient.

In order to clearly illustrate that this is not the case, we provide here a structured simulation with a data structure that resembles the one in TCB18-GAS in terms of observations, model specification, collinearity and efficiency level. The main difference here is that we simulate data with a given "true" efficiency that can be validated.

As we will show, the data structure using a simple linear cost function (as in TCB18) will show some negative regression coefficients due to collinearity. This has no impact on the choice of parameters and the model results converges quickly in DEA. However, the model does not yield reliable results in SFA, even for a time series.

We can therefore by construction dismiss the general claims in Oxera (2020): the regression signs for a larger model does not jeopardize the cost causality, it may occur for strictly positive cost functions, and it does not affect the correctness or convergence of the DEA scores.

4.1 Objectives

In this note we provide answers to the following relevant questions with respect to this type of model:

1) *Is there a negative trade-off between outputs in the actual data used in the DEA model? If so, to which extent? What role does the correlation between outputs play?*
2) *Can a negative trade-off between outputs be a problem in DEA? In our specific DEA model, given that we use free (i.e., strong) disposability as an assumption?*
3) *Did the fact that some of the outputs are correlated lead to any noteworthy effects in the TCB18 DEA model?*

The approach chosen here is to simulate data material resembling the TCB18 data, but with a given true cost function and efficiency level. In this manner we can abstract from the endogeneity of demonstrating fit for purpose with properties of the data used for its estimation, drawing upon the disputed assumptions.

4.2 Negative coefficients in regression results

The coefficients in a multiple regression model estimate the marginal change in the objective function value when changing an independent variable. In a general case and for a single variable, the sign indicates the direction of the change, such as the classification of (cost)-increasing and (cost)-decreasing parameters.

However, for larger models with a high level of collinearity among the parameters, the separation in sign no longer holds irrespective of the "true" causality among the parameters. In this case, one or several coefficients shift sign as the underlying independent variables co-vary and the nature of the causality of the parameters cannot be directly induced from the sign.

We differ between two different characterizations: negative cost-causality and substitution effects ("negative trade-off effects"). In the first case, an output parameter y_k is found to decrease the objective function (here: cost). This means that irrespective of the level of other parameters chosen, the cost function value decreases when y_k is increased. We can therefore say that y_k is not a "costdriver" but instead a favorable condition lowering the input.

In the second case, two output parameters, say y_i and y_j have a relationship such that an increase/decrease in one tends to imply the opposite effect for the other. This can be found as a negative correlation between the two variables and different signs for the regression coefficients in a given model. In this case we observe a substitution pattern, potentially but not necessarily linked to a cost causality. This could be a spurious (random) correlation or a choice of output profile that is not linked to cost.

The two cases are not equivalent to the case where an estimator simply has a negative sign because it corrects for another correlated factor already included in the model. E.g., in the electricity distribution models when including a general aggregate variable for decentralized power, adding a second parameter for a specific technology, say photovoltaics, comes out with a negative sign as the (cost increasing) factor is included in the first parameter and the second acts as correction (assuming that the impact is lower than the average in the aggregate factor).

In TCB18-GAS, we have no examples of cost-decreasing output parameters, they are all both statistically and techno-economically validated as cost-driving outputs. However, we have two parameters that have negative signs in a regression model for cost, indicating a high collinearity.

Thus, in order to model this specific dataset, we need to start with the strictly compliant positive cost function, still exhibiting the statistical characteristics of the TCB18-GAS sample. This will be done in the next section.

5. Simulated benchmarking data

In this section we describe how to simulate a structured data set with a number of units in a cross section, a large panel and four partially correlated outputs in a cost-function approach. Since we use simulated data, we change the names of the outputs not to confuse with the actual TCB18 model and its data, figuring here as a rail transport application.

5.1 Setup

We consider a set of 17 DMUs (observations or Decision Making Units, like TSOs in TCB18) over 10 years in a balanced panel, so where the panel has 170 observations.

There are five non-negative integer outputs; `y_railkm`, `y_switches`, `y_tonkm`, `y_passengers`, `y_freightvol` with mean and standard deviation as below. All DMUs share a common (unknown) cost function $C(y)$:

```
C = 100*y_railkm + 200*y_switches + 500*y_tonkm + 10*y_passengers + 10*y_freightvol + eps
```

where eps is a stochastic error. Some descriptive statistics for the dataset are given in Table 5-1 below. The observations for each DMU are stochastically independent, the distribution for cost and efficiency are stationary over time. However, the output parameters are internally correlated, see Table 5.2.

Table 5-1 Input and output variables, means and standard deviation (n=170).

Parameter	mean	St.dev
xCost	9,186,971	5,308,641
y_railkm	10,462	5,801
y_switches	9,927	5,987
y_tonkm	10,020	5,720
y_passengers	7,210	4,371
y_freightvol	2,131	1,501

Table 5-2 Correlation among inputs and outputs (full sample, n=170).

	xCost	y_railkm	y_switches	y_tonkm	y_passengers	y_freightvol
xCost	1.0000	0.9124	0.8660	0.9351	0.7126	0.3437
y_railkm	0.9124	1.0000	0.9122	0.9109	0.8268	0.3817
y_switches	0.8660	0.9122	1.0000	0.8261	0.7657	0.3408
y_tonkm	0.9351	0.9109	0.8261	1.0000	0.7617	0.3892
y_passengers	0.7126	0.8268	0.7657	0.7617	1.0000	0.3345
y_freightvol	0.3437	0.3817	0.3408	0.3892	0.3345	1.0000

The true efficiency in the sample is deterministic, stationary and given as below, acting on the single input (xCost).

```
xCost = C/Theta
```

Table 5-3 DMU and true efficiency (Theta).

DMU	Theta
1	0.7026
2	0.8772
3	0.6100
4	0.7763
5	0.9486
6	0.6289
7	0.9326
8	0.8780
9	1.0000
10	1.0000
11	1.0000
12	1.0000
13	1.0000
14	1.0000
15	1.0000
16	1.0000
17	1.0000
Mean	0.9032

We have thus a sample of 10 x 17 yearly observations of a strictly linear cost function with positive coefficients and a known distribution for inefficiency with distribution similar to that of TCB18.

5.2 Single factor regressions

In Table 5-4 the single-parameter regression results for the full panel are presented for each parameter candidate. It is shown that the five parameters separately are significant with or without a constant to explain the dependent variable (xCost). However, the level of significance highly differs between the two prior (y_railkm and y_tonkm, both adj $R^2 > 0.83$) versus the least significant (y_freightvol, adj $R^2 < 0.45$).

Table 5-4 OLS results for single-parameter models with intercept, n=170.

	Dependent variable:				
	xCost				
	(1)	(2)	(3)	(4)	(5)
y_railkm	835.102*** (28.863)				
y_tonkm		851.828*** (28.419)			
y_switches			767.416*** (34.275)		
y_passengers				962.856*** (57.101)	
y_freightvol					2,360.032*** (203.337)
Constant	450,441.300 (345,032.300)	651,875.000** (327,644.600)	1,568,918.000*** (397,021.500)	2,245,060.000*** (481,059.400)	4,158,701.000*** (529,374.600)
Observations	170	170	170	170	170
R^2	0.833	0.842	0.749	0.629	0.445
Adjusted R^2	0.832	0.842	0.747	0.626	0.442
Residual Std. Error (df = 168)	2,176,765.000	2,113,269.000	2,667,567.000	3,244,857.000	3,966,549.000
F Statistic (df = 1; 168)	837.149***	898.458***	501.303***	284.337***	134.711***

Note: *p<0.1; **p<0.05; ***p<0.01

The parameter plots towards cost are given in Figure throughout Figure . The plots confirm the observations in Table 5-4, the parameters have visually a monotonic increasing trend with varying level of variability.

The single-parameter results for the cross sections (single year) are provided below they show the same qualitative results with lower explanatory power.

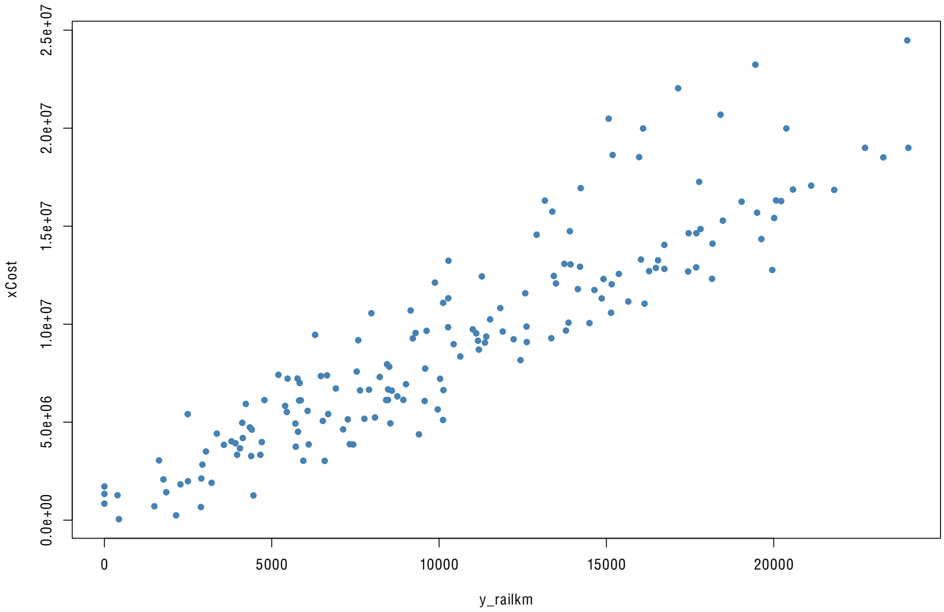

Figure 1 y_railkm vs xCost, n=170.

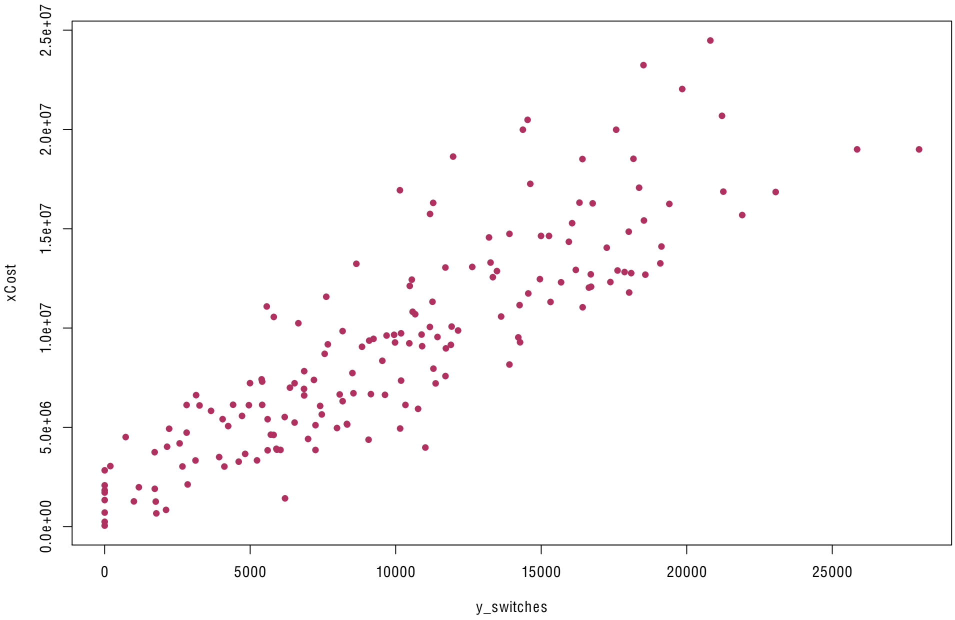

Figure 2 y_switches vs xCost, n = 170.

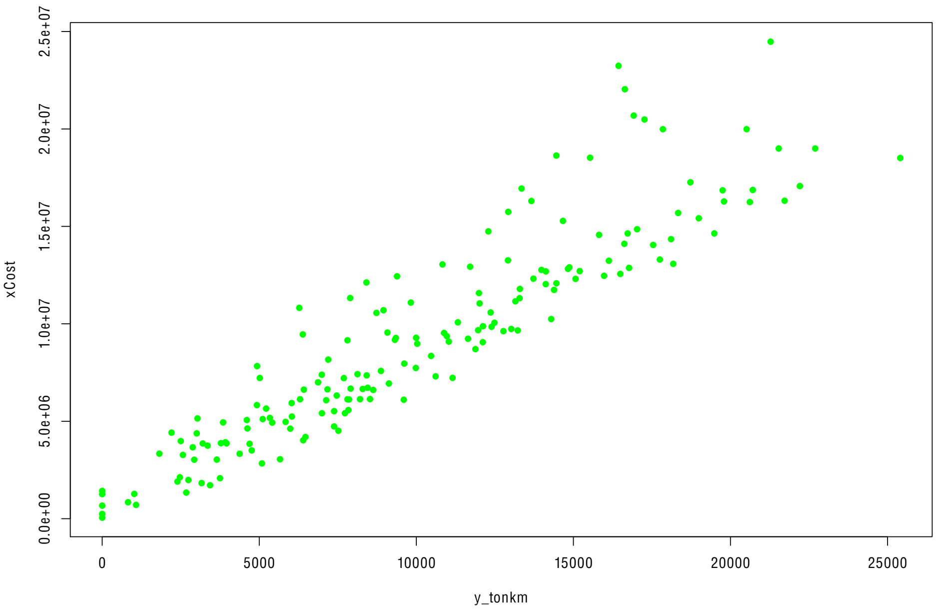

Figure 3 y_tonkm vs xCost, n =170

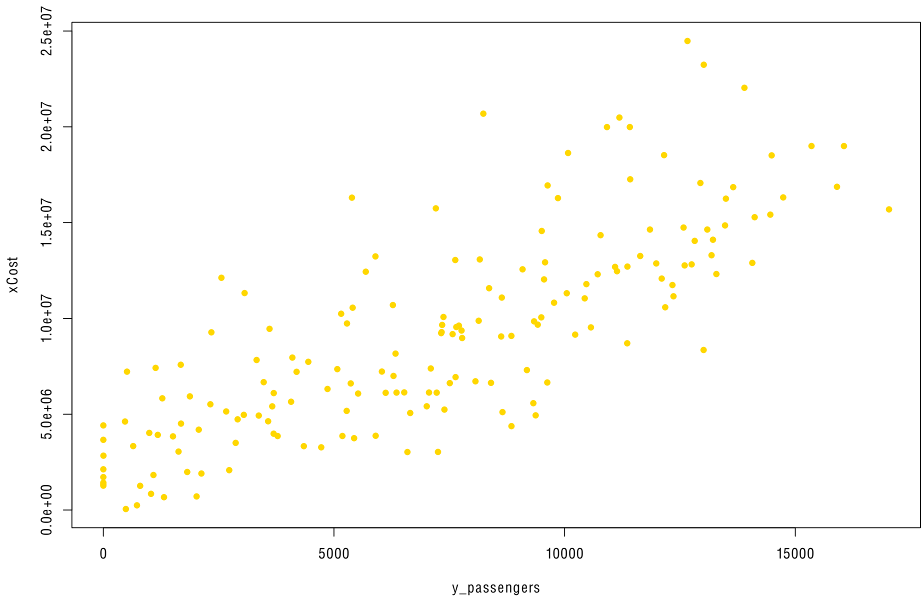

Figure 4 y_passengers vs xCost, n =170.

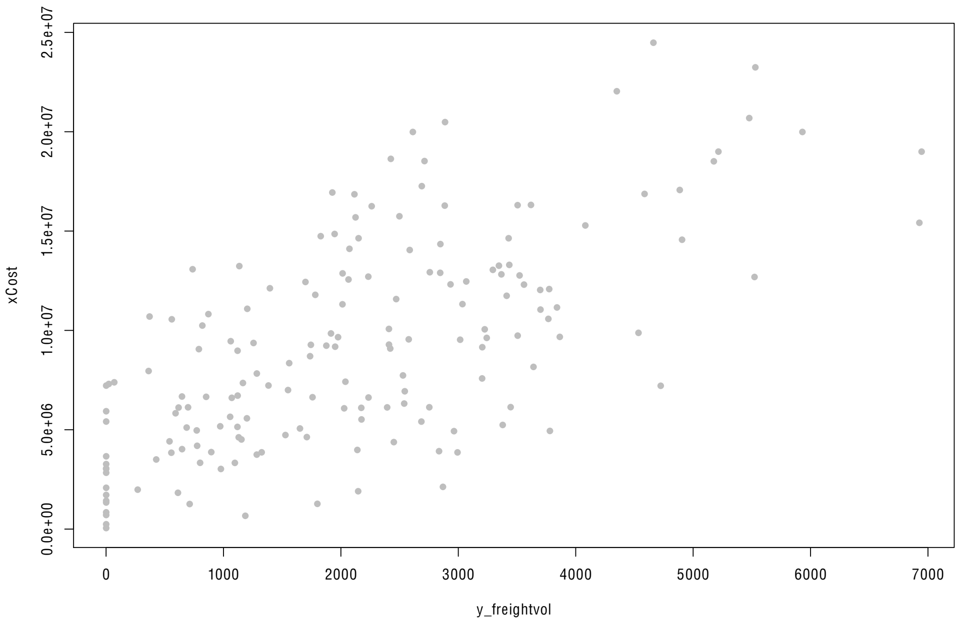

Figure 5 y_freightvol vs xCost, n = 170.

5.3 Model specification

In a model specification effort, we combine the parameters in Table 5-5. As confirmed by an ANOVA or AIC analysis, adding up to four parameters increases the explanatory power of the model even when considering the model size.

However, we note that the full model with five parameters (indeed the true model) does not come out as preferred, since the parameter y_freightvol loses its significance in any model with more than three parameters. It is therefore interesting to imagine that the model specification at this stage would consider halting at a model with four parameters.

We also note that the coefficient for the parameter y_passengers shifts to negative (significant) in the preferred four-parameter model. This is due to the collinearity among the parameters. In Table 5-6 we show the Belsley condition index (CI) and the variance decomposition shares for the four parameters. A proposed cut-off level for the condition index is 30 (Belsley et al, 1980), which means that no parameter is a priori compromised. The variance decomposition proportions show how the variance in the determination of a specific coefficient for one variable relates to the presence of another parameter. For instance, the y_railkm parameter is virtually absorbing the significance of the constant when present (share 0.972) (which in fact is correct – there is no fixed term in the true model).

As with regard to y_passenger, the parameter also has a strong variance interaction with y_railkm (not to be confounded with the correlation in Table 5-2). This means here that the joint determination of the two coefficients is not reliable, there is collinearity at hand.

Table 5-5 OLS results, 2 to 5 parameter models (n=170).

	Dependent variable:			
	xCost			
	(1)	(2)	(3)	(4)
y_railkm	406.580*** (64.121)	235.787*** (88.433)	341.173*** (97.526)	338.871*** (101.306)
y_tonkm	472.741*** (65.031)	482.727*** (63.896)	494.492*** (63.178)	494.720*** (63.423)
y_switches		170.064*** (61.896)	170.843*** (61.019)	170.295*** (61.523)
y_passengers			−173.653** (72.037)	−173.168** (72.467)
y_freightvol				12.171 (139.287)
Constant	196,733.100 (303,629.900)	195,244.800 (297,846.200)	219,099.800 (293,788.400)	216,920.400 (295,729.500)
Observations	170	170	170	170
R^2	0.873	0.879	0.883	0.883
Adjusted R^2	0.872	0.876	0.880	0.879
Residual Std. Error	1,902,866.000 (df = 167)	1,866,616.000 (df = 166)	1,840,141.000 (df = 165)	1,845,700.000 (df = 164)
F Statistic	574.169*** (df = 2; 167)	400.307*** (df = 3; 166)	310.384*** (df = 4; 165)	246.816*** (df = 5; 164)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5-6 Belsley condition index and variance decomposition shares, n=170.

	Constant	y_railkm	y_switches	y_tonkm	y_passengers
Constant	0.009	0.001	0.002	0.002	0.002
y_railkm	0.972	0.003	0.012	0.006	0.015
y_switches	0	0.003	0.31	0.004	0.691
y_tonkm	0.013	0.001	0.248	0.61	0.133
y_passengers	0.006	0.992	0.428	0.379	0.159
Condition index	1.00	4.92	9.92	10.85	20.75

We investigate alternative 3-parameter models in Table 5-7 to explore whether the parameter y_passengers is necessary and useful. As seen, the coefficient for y_passengers is negative in all models, but not significant unless y_railkm is in the model. The preferred model using a criterion such as BIC or Mallow's Cp is the full four-parameter model that we call Model 4.

If the non-significant intercept is removed from the estimation of Model 4, the resulting model in Table 5-8 is obtained. As seen, the adjusted fit now attains a level of 97%, with coefficient values and signs similar to the ones with intercept.

Table 5-7 OLS results for alternative 3 and 4-parameter models, n=170.

	Dependent variable:			
	xCost			
	(1)	(2)	(3)	(4)
y_railkm	341.173*** (97.526)		512.096*** (77.607)	786.062*** (92.522)
y_tonkm	494.492*** (63.178)	623.306*** (53.048)	484.388*** (64.362)	
y_switches	170.843*** (61.019)	304.460*** (49.170)		143.564** (71.122)
y_passengers	−173.653** (72.037)	−60.687 (66.538)	−172.584** (73.505)	−130.097 (83.851)
Constant	219,099.800 (293,788.400)	356,798.800 (300,833.100)	220,448.000 (299,779.000)	476,297.700 (340,840.600)
Observations	170	170	170	170
R^2	0.883	0.874	0.877	0.839
Adjusted R^2	0.880	0.872	0.875	0.836
Residual Std. Error	1,840,141.000 (df = 165)	1,901,409.000 (df = 166)	1,877,665.000 (df = 166)	2,148,333.000 (df = 166)
F Statistic	310.384*** (df = 4; 165)	383.785*** (df = 3; 166)	394.960*** (df = 3; 166)	288.643*** (df = 3; 166)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5-8 OLS result for model 4 (without intercept), n=170.

	Dependent variable:
	xCost
y_railkm	350.918***
	(96.517)
y_switches	170.918***
	(60.937)
y_tonkm	499.762***
	(62.698)
y_passengers	−171.843**
	(71.900)
Observations	170
R^2	0.971
Adjusted R^2	0.970
Residual Std. Error	1,837,680.000 (df = 166)
F Statistic	1,373.247*** (df = 4; 166)
Note:	*p<0.1; **p<0.05; ***p<0.01

5.4 DEA estimations

In this section, we investigate questions (2) and (3) above, i.e., whether a DEA model based on collinear outputs would yield biased results. In the given example, we know the "true" score as the distribution is generated from a controlled and pre-defined cost function. A simple way of examining the bias issue is then to consider the ratio of the estimated DEA score to this true score, following Mehdiloo and Podinovski (2019). If no bias is at hand, the ratio is 100%, if the score is lower it means that the DEA model underestimates the efficiency leading to a bias.

We now run two potential models in DEA framework (constant returns to scale, no outlier detection, no scaling), see Table 5-9 below.

Model 4 is based on the OLS estimations above and could be a likely candidate for a model, deleting the non-significant parameter y_freightvol from the model. Model 5 is the true model for reference.

Table 5-9 Benchmarking models 4 and 5.

Model 5	Model 4
xCost	xCost
y_railkm	y_railkm
y_switches	y_switches
y_tonkm	y_tonkm
y_passengers	y_passengers
y_freightvol	

The DEA models are run in two modes: annually ("annual") with as reference set only the units for the given year, and pooled ("panel") with all units included in the reference set. The mean results are presented in Table 5-10 below along with the true efficiency scores.

The results reveal that the mean efficiency scores are close to the true values, both for model 4 and 5. Naturally, model 5 gives an equal or higher score compared to model 4 (with one output missing), but the difference is small. As expected, the panel estimates are very close to the true values with model 5 slightly higher than model 4.

A closer look at the differences compared to the true values by DMU are given in Table 5-11. The color indicates the sign of the difference, green meaning that the DEA estimate is more generous than the true value. Here the results nicely illustrate the minimum extrapolation feature of DEA, the annual results for both model 4 and 5 are more conservative than the true values, up to 7.2% higher than the actual score. Note also that the models correctly identify all real peers within a 0.4% range.

In terms of ranges for the individual DMU, the results are presented in Table 5-13.As before, the cautiousness is clearly shown with occurrences of full efficiency for inefficient units in some periods. The differences with respect to the true values are illustrated in Table 5-14. As seen, the correct model 5 never underestimates the real efficiency, but frequently overestimates it. Model 4 slightly penalizes the fully efficient peers, but with less than 0.5%. However, model 4 is also somewhat less generous in the overestimation than model 5, an effect of lower dimensionality and the variability of the 5th parameter.

Thus, returning to the initial questions (2) and (3), we see in Table 5-12 that the bias from the full model in DEA is positive 2.6% and a smaller model gives 1.7%. The bias is not linked to the model specification but to the effect of individual variations in cross-section estimations with 17 DMU. For a panel application, the bias is virtually eradicated (+0.5% for the full model and -0.1% for the smaller model). We may therefore safely conclude that not only DEA correctly identifies the peers for a model with collinear outputs, but even with a limited sample size it limits the bias by adding a cautious overestimation thanks to the minimum extrapolation principle, of 1.7%-2.6% for small cross-sections. More complex assumptions for the production space, such as the hybrid disposable technology in Mehdiloo and Podinovski (2019), can therefore not offer any real improvement in the estimation, as the sample size effects are independent of the estimation technology.

Table 5-10 True and estimated mean efficiency scores, models 4 and 5.

TSO	True Theta	DEA(5, annual)	DEA(5, panel)	DEA(4, annual)	DEA(4, panel)
1	0.946	0.964	0.951	0.964	0.950
2	0.644	0.716	0.646	0.715	0.645
3	0.710	0.763	0.710	0.732	0.708
4	0.998	1.000	0.999	1.000	0.996
5	0.911	0.944	0.922	0.938	0.918
6	0.605	0.659	0.641	0.622	0.603
7	0.823	0.883	0.826	0.859	0.821
8	0.768	0.794	0.768	0.789	0.766
9	1.000	1.000	1.000	0.999	0.996
10	1.000	1.000	1.000	1.000	0.998
11	1.000	1.000	1.000	1.000	0.997
12	1.000	1.000	1.000	1.000	0.999
13	1.000	1.000	1.000	0.999	0.998
14	1.000	1.000	1.000	1.000	0.998
15	1.000	1.000	1.000	0.999	0.998
16	1.000	1.000	1.000	0.999	0.997
17	1.000	1.000	1.000	1.000	0.998
Mean	**0.906**	**0.925**	**0.910**	**0.919**	**0.905**

Table 5-11 Difference in mean (absolute values) from true efficiency.

	Differences			
TSO	DEA(5, annual)	DEA(5, panel)	DEA(4, annual)	DEA(4, panel)
1	0.018	0.005	0.018	0.004
2	0.072	0.002	0.071	0.001
3	0.053	0	0.022	-0.002
4	0.002	0.001	0.002	-0.002
5	0.033	0.011	0.027	0.007
6	0.054	0.036	0.017	-0.002
7	0.06	0.003	0.036	-0.002
8	0.026	0	0.021	-0.002
9	0	0	-0.001	-0.004
10	0	0	0	-0.002
11	0	0	0	-0.003
12	0	0	0	-0.001
13	0	0	-0.001	-0.002
14	0	0	0	-0.002
15	0	0	-0.001	-0.002
16	0	0	-0.001	-0.003
17	0	0	0	-0.002
Mean	**0.019**	**0.004**	**0.013**	**-0.001**

Table 5-12 Misspecifcation bias for DEA scores per DMU.

	Misspecification bias			
TSO	DEA(5, annua	DEA(5, panel	DEA(4, annua	DEA(4, panel)
1	1.019	1.005	1.019	1.004
2	1.112	1.003	1.110	1.002
3	1.075	1.000	1.031	0.997
4	1.002	1.001	1.002	0.998
5	1.036	1.012	1.030	1.008
6	1.089	1.060	1.028	0.997
7	1.073	1.004	1.044	0.998
8	1.034	1.000	1.027	0.997
9	1.000	1.000	0.999	0.996
10	1.000	1.000	1.000	0.998
11	1.000	1.000	1.000	0.997
12	1.000	1.000	1.000	0.999
13	1.000	1.000	0.999	0.998
14	1.000	1.000	1.000	0.998
15	1.000	1.000	0.999	0.998
16	1.000	1.000	0.999	0.997
17	1.000	1.000	1.000	0.998
Mean	**1.026**	**1.005**	**1.017**	**0.999**

Table 5-13 Ranges (min-max) for annual DEA scores (10 years) per DMU.

	DEA(5, annual)		DEA(4, annual)	
TSO	min	max	min	max
1	0.947	1.000	0.947	1.000
2	0.644	1.000	0.644	1.000
3	0.710	0.980	0.710	0.818
4	0.999	1.000	0.999	1.000
5	0.911	1.000	0.911	1.000
6	0.605	1.000	0.604	0.712
7	0.823	1.000	0.820	1.000
8	0.769	0.863	0.766	0.863
9	1.000	1.000	0.997	1.000
10	1.000	1.000	0.999	1.000
11	1.000	1.000	0.998	1.000
12	1.000	1.000	0.999	1.000
13	1.000	1.000	0.998	1.000
14	1.000	1.000	1.000	1.000
15	1.000	1.000	0.998	1.000
16	1.000	1.000	0.997	1.000
17	1.000	1.000	0.998	1.000

Table 5-14 Differences from true values (max, min) for model 4 and 5 (annual).

	DEA(5, annual)		DEA(4, annual)	
TSO	min	max	min	max
1	0.001	0.054	0.001	0.054
2	0.000	0.356	- 0.000	0.356
3	- 0.000	0.270	- 0.000	0.108
4	0.001	0.002	0.001	0.002
5	0.000	0.089	- 0.000	0.089
6	- 0.000	0.395	- 0.001	0.107
7	0.000	0.177	- 0.003	0.177
8	0.001	0.095	- 0.002	0.095
9	- 0.000	-	- 0.003	-
10	-	-	- 0.001	-
11	-	-	- 0.002	-
12	-	-	- 0.001	-
13	-	-	- 0.002	-
14	-	-	-	-
15	-	-	- 0.002	-
16	- 0.000	-	- 0.003	-
17	-	-	- 0.002	-

5.5 SFA estimations

Stochastics Frontier Analysis (SFA, see Aigner et al., 1977) is a technique that estimates both general noise in the production or cost function, as well as a firm-specific inefficiency term. The primary assumption behind SFA is that the production function is stochastic, i.e. the outcome in terms of yield or cost depends on some randomness. Although conceptually very appealing and indispensable when dealing with applications such as e.g. energy generation and agriculture, SFA contrary to DEA relies on a series of technical assumption with regard to the distribution of the stochastic noise and the efficiency. These assumptions are not easily verifiable (e.g. is the distribution of inefficiency distributed as a half-normal, gamma, beta or as a truncated normal?), nor harmless (direct impact on the level and distribution of scores).

Most regulatory benchmarking relies on predominantly deterministic models as tariff setting and review should not depend on random effects or their estimation. For this reason, SFA is rarely used as primary model in regulation, but more often as cross-validation instrument.

SFA is computationally considerably heavier than e.g. linear regression, forming a nonlinear optimization problem for which no closed form solution exists. Thus, iterative numerical procedures are used to estimate the maximum likelihood for the model. These models do not always converge, in particular for certain types of formulations.

A model without any transformation of the input-output data is called a 'level-model'. In DEA this corresponds to some strong properties since the reaction in e.g. cost for increasing/decreasing one unit of an output is determined by the data (the frontier) and not by any assumption.

A model such as log(Totex) = log(Lines) + log(Power) is called a log-log model. Here the cost consequence of adding one unit of lines is defined by the log, although the data may say otherwise locally. A more advanced model may have interaction terms such as log(Totex) = log(Lines) + log(Power) + log(Lines*Power), called a translog function.

The SFA in general does not perform well with level data, it converges best with relatively small models on normalized data, e.g. log-linear formulations. Model 4 in levels (non-transformed data) can therefore not be used to estimate meaningful efficiencies in SFA. Obviously, this does not mean that the inefficiency disappears in SFA, it is a consequence of the high fit of the model not leaving enough variance to allow for the estimation of both an independent noise and half-normal inefficiency term.

A transformation to a log-model works in the Battese and Coelli (1992) implementation on R (frontier, version 1.1-8, by Henningsen) with the results below, compared to DEA and the true efficiency level. As seen in Figure 3 below, the DEA results quickly converge to the true (green) level whereas in this case the SFA results (model 4, log) converge to a lower level with a higher variance. The initial estimates for SFA are highly variable. Note that as the sample size decreases in SFA, the model cannot separate noise from inefficiency, leading to very high efficiency estimates, finally becoming non-significant. It is this effect we note at the outset, where initially SFA is higher and then decreasing. Had we tried with smaller sample sizes, then SFA would not have returned any significant estimates, which might have been misinterpreted as if there were no inefficiency in the data.

Below we also present the results for SFA models for a cross-section and for the largest instance (pooled with n=170) with level and log formulations. Notice that y_passengers is significant and with a negative sign in the cross-section result but turning to non-significant in the full pooled instance for the log-formulation. In the level model, the sign for y_passengers switches to positive, but at the cost of the identification of inefficiency. Thus, as stated above, SFA is not an adequate and reliable estimation technology for all types of cost functions, including the level-based function with collinear outputs as in TCB18 and the example.

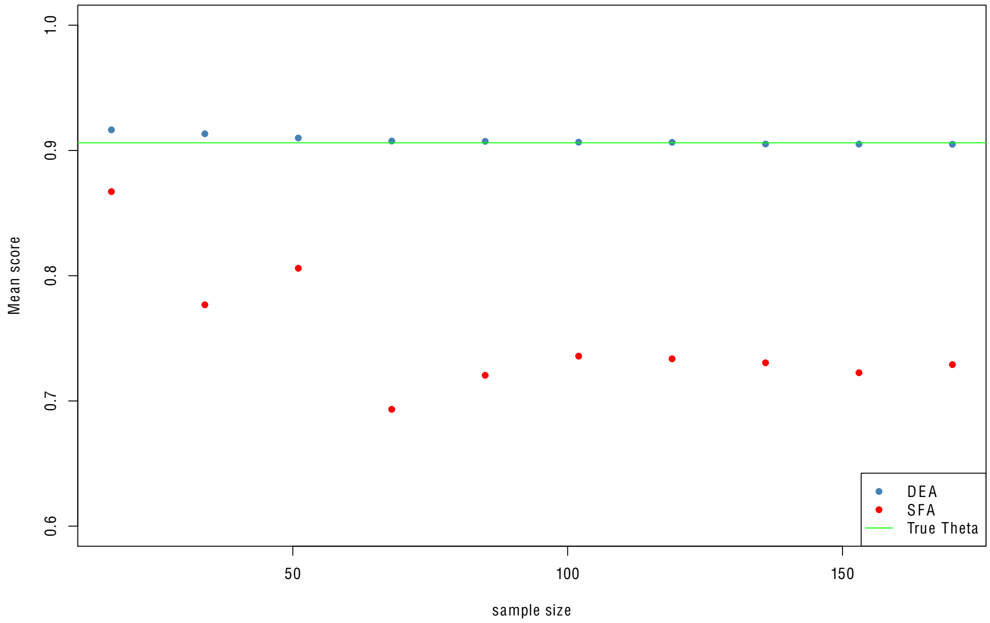

Figure 3 Mean SFA, DEA and true efficiency scores vs sample size.

SFA results (cross section 10, log, n=17)

```
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)

final maximum likelihood estimates
                         Estimate Std. Error  z value Pr(>|z|)
(Intercept)             6.8274106  0.7670439   8.9009  < 2e-16 ***
log(y_railkm + 1)       0.1553354  0.2966993   0.5235  0.60060
log(y_tonkm + 1)        0.7223927  0.2637299   2.7391  0.00616 **
log(y_switches + 1)     0.1847101  0.0727741   2.5381  0.01114 *
log(y_passengers + 1) -0.0797643  0.0065106 -12.2514  < 2e-16 ***
sigmaSq                 0.0279558  0.0123505   2.2635  0.02360 *
gamma                   1.0000000  0.0068761 145.4313  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 15.00525

cross-sectional data
total number of observations = 17

mean efficiency: 0.8902327
```

SFA results (pooled, log, n=170)

```
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)

final maximum likelihood estimates
                       Estimate Std. Error z value  Pr(>|z|)
(Intercept)           9.2343238  0.2488652 37.1057 < 2.2e-16 ***
log(y_railkm + 1)     0.1909880  0.0376984  5.0662 4.058e-07 ***
log(y_tonkm + 1)      0.3705953  0.0224085 16.5382 < 2.2e-16 ***
log(y_switches + 1)   0.1351792  0.0202158  6.6868 2.281e-11 ***
log(y_passengers + 1) 0.0070172  0.0141768  0.4950    0.6206
sigmaSq               0.2428913  0.0328299  7.3985 1.378e-13 ***
gamma                 0.9393264  0.0281304 33.3919 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: -31.64057

cross-sectional data
total number of observations = 170

mean efficiency: 0.7290059
```

SFA results (pooled, level, n=170)

```
Error Components Frontier (see Battese & Coelli 1992)
Inefficiency increases the endogenous variable (as in a cost function)
The dependent variable is logged

final maximum likelihood estimates
               Estimate Std. Error    z value  Pr(>|z|)
y_railkm     1.0587e+02 1.3331e+00 7.9418e+01 < 2.2e-16 ***
y_switches   1.9792e+02 6.7594e-01 2.9280e+02 < 2.2e-16 ***
y_tonkm      4.9625e+02 6.4158e-01 7.7349e+02 < 2.2e-16 ***
y_passengers 6.6792e+00 1.0817e+00 6.1745e+00 6.636e-10 ***
sigmaSq      8.4805e+12 1.0000e+00 8.4805e+12 < 2.2e-16 ***
gamma        1.0000e+00 2.5161e-08 3.9744e+07 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: -2617.615

cross-sectional data
total number of observations = 170

mean efficiency: NaN
```

6. Conclusions

The simulation in this section is relevant to TCB18 for the following reasons:

1. Model size and number or observations are similar
2. The correlation among parameters is similar
3. The variable decomposition factors and conditions index are similar
4. The coefficients, signs and shift of sign are corresponding to model size.
5. The single-variable results and plots are equivalent.

For the given example we show for a random dataset with a random inefficiency that:

1. The sign of a particular parameter does not affect the estimation in DEA or SFA.
2. That a DEA estimation using parameters having negative coefficients in some OLS obtains correct results.
3. That DEA estimates asymptotically converge to the true level of efficiency from above (cautiousness).
4. That DEA estimates are relatively insensitive to the model specification for collinear parameters.
5. That an SFA estimation for a standard linear untransformed model may not result in any useful estimation even for a controlled experiment
6. That an SFA estimation for a loglinear model may result in scores that are lower than the true value and that convergence is slow.

In this appendix we have provided answers to the following relevant questions:

1) Is there negative trade-off between outputs in the actual data that we use in the DEA model? And if so, to what extent? What role does correlation between outputs play?

    No, there are no negative tradeoffs among the outputs. There is collinearity among the parameters leading to estimates of the coefficients and their signs that limit their direct use. The OLS model for average cost can only be used to predict the level of cost, not to evaluate the tradeoffs among the parameters.

2) Can a negative trade-off between outputs be a problem in DEA? And in our specific DEA model, also given that we use free (i.e. strong) disposability as an assumption?

    In general, the use of an input in place of an output in a DEA model leads to erroneous results. However, this is not the case in TCB18 where all outputs are validated both statistically and techno-economically. See the example for a constructive proof.

3) Did the fact that some of the outputs are to an extent correlated lead to any noteworthy effects in the TCB18 DEA model?

    No, DEA is very robust to correlated outputs, both in practice and in theory. However, SFA reacts poorly to both the collinearity and the type of model (level, linear).

Thus, we can conclude the following:

> The estimation of cost functions for collinear parameters does not affect the overall precision of OLS predictions, nor of DEA estimates of efficiency. However, collinearity in the model specification limits the use of certain tools for model validation.

SUMICSID SPRL (BELGIUM)
Rue Maurice Lietart 56
B-1150 Bruxelles, BELGIUM
belgium @ sumicsid.com

Tel: +32 26 75 15 23
Fax: +32 27 06 53 58

SUMICSID APS (DENMARK)
Fru Ingesvej 19
DK-4180 Sorø, DENMARK
denmark @ sumicsid.com

Tel +45 57 83 15 18
Fax: +32 27 06 53 58

http://www.sumicsid.com