Appraising the benefits of bottleneck removal in rail transport: a simplified CBA approach

Beria, Paolo and Grimaldi, Raffaele

Politecnico di Milano , Politecnico di Milano

10 May 2013

POLITECNICO DI MILANO

LABORATORIO DI POLITICA DEI TRASPORTI
TRASPOL
RESEARCH CENTER ON TRANSPORT POLICY

Appraising the benefits of bottleneck removal in rail transport: a simplified CBA approach.

Paolo Beria, Raffaele Grimaldi
Politecnico di Milano – Dipartimento di Architettura e Studi Urbani.

May 10th, 2013

Abstract

The removal of infrastructure bottlenecks is widely considered among the most profitable interventions, in socio-economic terms, and rail transport is not an exception. However, as outlined for example by RailPAG (2005), the measurement of the related benefits is difficult and no specific manuals indications seem to exist. From a general point of view, by removing a rail bottleneck we expect at least two kinds of benefits: direct benefits to transport users and external benefits to the rest of society (environmental externalities, accidents and congestion) due to the avoidance of possible shift to more impactful transport modes. The first effect is particularly hard to correctly evaluate, especially without a complete transport model, and thus CBAs currently performed might often result biased.

The aim of this paper is to propose a simplified approach to estimate the effects of a capacity constraints for a simple rail network, and assess its removal through a CBA.

In the first part, we briefly analyse the transport economics literature on the issue. In the following we introduce the proposed methodology, based on the use of a standard logit model, to measure the rail users' generalised costs with and without the capacity constraint, and the consequent users and social surplus variation. The model is specified initially for a single link and then extended to a more complex network. Then, we outline the other elements to be included into a CBA in addition to surplus variation: rail service performance improvements, external costs associated to road shift and possible wider economic effects. We also discuss the effect of regulation in the distribution of calculated surplus variations.

1. Introduction

This working paper includes a preliminary formulation of an approach to the calculation of bottlenecks removal benefits, to be consistently used in a Cost Benefit Analysis (CBA).

The “simplification” consists in the fact that the methodology is suitable to be applied without a full transport simulation model describing the upstream and downstream network and the complete origin-destination (OD) matrix of traffic. Instead, the methodology is based on two components: a calibrated logit model - capable to simulate the modal shares of traffic in the bottlenecked section only - and the so called “rule of half”, to consistently calculate the user surplus variations in a CBA without explicitly knowing the demand function.

The paper is structured as follows. We initially introduce (Section 2) how known literature deals with the benefits of bottleneck removal into CBAs. Section 3 describes the core of the methodology, applied to the simplest case of single segment network. The following Section 4 extends the base case to a generic network, in which one or more segments may reach saturation, proposing a simple algorithm to calculate the corresponding extra-costs and traffic. Section 5 briefly outlines a parallel but crucial aspect: how market conditions and capacity allocation mechanisms influence the users and the social costs of bottleneck. Finally, in Section 6 we describe how a full CBA, including also the shadow cost of bottleneck, must be performed. Section 7 concludes.

2. The core benefits of bottleneck removal: scarcity and congestion

The way a bottleneck removal should be assessed in socio-economic CBAs seems quite neglected in the literature, quite surprisingly considering its importance and representing a very common issue. The *Railway Project Appraisal Guidance* (RailPAG: EC-EIB, 2004) stated that *“the measurement of benefits from bottleneck removal is obviously very difficult. In particular because, under congestion conditions, there are always substantial tradeoffs (for instance between additional traffic and safety conditions) that are difficult to estimate. This is, in any case, a technical question that requires substantial research and should be dealt with in specific manuals.”*

The same guidance proposes a case study example about the appraisal of a bottleneck removal (Case Study 3, page 96), however no specific methodology is proposed to estimate the benefits: *“A bottleneck will typically be associated with congestion. Small deviations to train plans (train delays) will result in the need to readjust train schedules significantly. A removal of the bottleneck will hence often have significant repercussions on reliability. The improvement in reliability will affect the costs of the rail operators, and these savings could be significant but difficult to trace and measure. Improvement in reliability will also lead to shorter travel times and so to an improvement in the welfare of travellers over and above average time savings. These improvements will have to be valued separately and accounted for in the analysis, otherwise the real benefits of the project will not be covered. On the other hand attention must be paid not to double-count time savings under the “reliability” label.”* The guidance focuses on effects for users remaining on the rail line, but does not consider what happens to the others.

In the Strategic Business Case of the British “New Lines Program ” (Network Rail, 2009), freight benefits due to a release of freight paths on the conventional West Coast Main Line after the building of the planned parallel High Speed 2 line, have been calculated using the so called “sensitive lorry miles avoided” approach (SRA, 2003): they attach to each kilometer (mile) that is expected to shift from road to rail transport after the bottleneck will be removed a net external cost saving (i.e. the difference between generated gross external costs – in terms of accidents, congestion and environmental impacts – and paid fuel taxes). Again, it seems that no surplus gain has been assessed for trains excluded from the congested rail line without the investment.

Nilsson (2012) explains how congestion, in the way it is usually understood in road transport, takes the form of two different components in rail transport:

- a scarcity problem; and,
- a congestion problem.

The problem of scarcity rises when it becomes difficult, and gradually impossible, to build a timetable capable of guarantee all the requested slots. The problem of congestion instead appears after the timetable has been designed, since the planned timetable might be disturbed by delays and irregularities.

With this respect, Nash and Samson (1999) outlined how *“the main consequence of full utilisation of capacity is that users simply cannot get the capacity they want when they want it; they have to run their trains at times and possibly speeds different to their preferred alternative, or to give up the journey.”*

Eliasson and Borjesson (2012) discuss about how rail travel time-capacity relationships, that would allow to estimate the benefits of a bottleneck removal, is a field that would require further research. The authors demonstrate how timetable assumptions influence the results of railway investment appraisals. In the Swedish appraisal practice (Eliasson and Borjesson, 2012) train travel time-capacity relations are provided by the official guidelines, and the most common approach is to simulate the same planned timetable in both the do nothing and investment alternatives and compare the resulting travel times.

Our approach instead tries to practically evaluate the socio-economic effects of bottleneck, starting from its consequence, that is an artificially constrained mode share, trying to simplifying and overcoming the difficulties related to the construction of such a relationship. We will focus in the following sections on the appraisal of the main consequence of rail bottlenecks: the scarcity problem.

In this paper we will not deal the issue of how rail capacity can be assessed and which values should be chosen: it is however clear that this is a crucial issue, as its value – and the consequent timetable assumptions (Eliasson and Borjesson, 2012) - determines how much scarcity or congestion will rise, thus changing the results of the CBA.

3. Base case: single line and single relationship

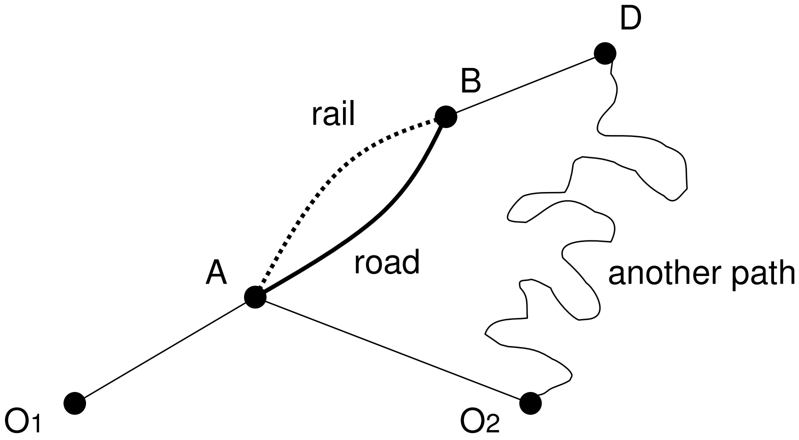
1.1 The model

Let’s consider a relation A to B, where two transport modes compete, for example “rail” and “road”.

We assume that the traffic and the modal split are known in year 0. We also assume that it is possible to make an estimation of average generalised costs for the users of the two modes passing through A-B. To the contrary, in absence of a full transport model, we cannot estimate the generalised cost (hereinafter “GC”) for every type of user and for every OD couple, neither on the currently used alternative nor on all the other possible alternatives. So, we cannot simulate their behaviour and moreover the rise in generalised costs due to bottleneck.

An example (Figure 1) can help us in understanding the point.

Figure 1 - Example of network



Users of the section A-B come from the two origins O_1 and O_2 and all go to the destination D. We do not exactly know their GCs, neither where they exactly come from. We just know that their paths O_1ABD and O_2ABD pass through section A-B, divided into the two available modes, “rail” and “road” (both available from O_1 to D and from O_2 to D). Actually, their choices depend on many individual characteristics, like their origin, their upstream and downstream paths and modes, their value of time (in its different component), their transport costs (both depending on what they carry), etc.

Let’s consider the case in which traffic increases and a bottleneck in the A-B “rail” segment occurs. In this case they will individually evaluate their best possible substitute options, choosing the best one. This is what transport models simulate, starting from a full OD matrix and from a full network description. For example, some of them (say those from O_1) could prefer to switch to the “road” mode, because the extra-cost for them is lower than the extra-cost of the saturated rail line and of any other known alternative. To the contrary, others (say those from O_2) could radically change path and now prefer the link “another path”.

The scarcity cost of the bottleneck (more precisely, its “unit shadow cost”) is then dependant on:

1. the OD matrix;
2. the network characteristics;
3. the users characteristics.

Many of these aspects cannot be fully understood without a transport model and, consequently, the benefit estimation of the bottleneck removal is hard to quantify.

1.2 Generalised costs and traffic under saturation

However, random utility models (typically logit) applied to a specific network section are used exactly to describe the behaviour of heterogeneous users starting from the average costs (or utilities; see Cascetta, 2001) and assuming a particular distribution of the singular costs (or utilities), not included in the deterministic measure.

So, through a logit model it is possible to estimate a modal share due to average conditions, given a certain distribution of individual behaviours, without knowing the behaviour of single users. In this example, we will refer to a simple binomial logit, like the one in Equation 1.

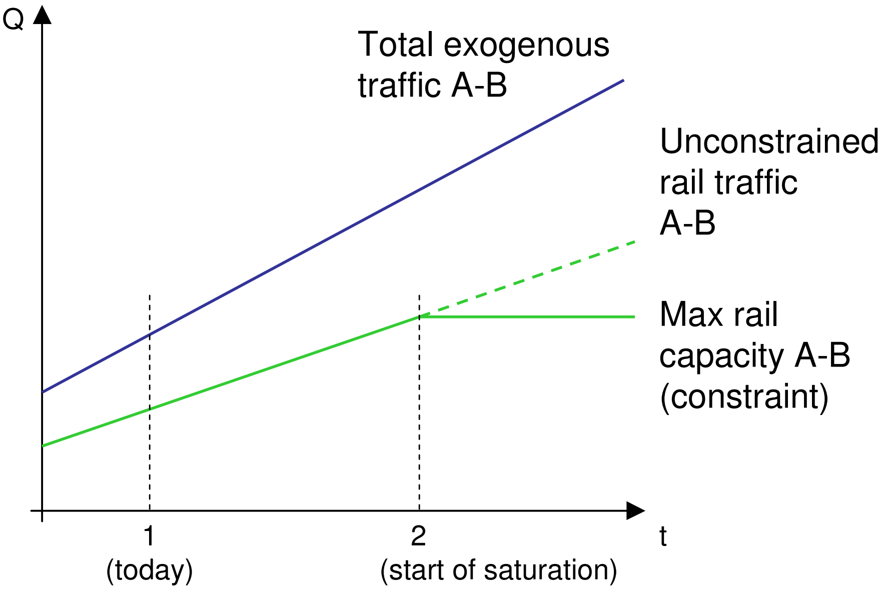
$$p_{rail} = \frac{\exp(\lambda \cdot GC_{rail})}{\exp(\lambda \cdot GC_{rail}) + \exp(\lambda \cdot GC_{road})}$$

Equation 1

Where p_{rail} is the mode share of rail transport, GC_{rail} and GC_{road} are the average generalised costs of respectively rail and road transport, and λ is the calibration parameter of the model.

The following Figure 2 depicts the concept.

Figure 2 - Exogenous and constrained traffic



Rail, road and the consequent total traffic on the section A-B would exogenously rise. However, at a certain moment, rail increase is capped as the section reaches its capacity: exogenous increase is no more possible and part of traffic shifts to its second best alternative, in this simplified case the parallel road.

In particular, we use logit models in two steps. The first is necessary to estimate the parameters (calibration), if not otherwise known.

Step 1: GC_{rail} , GC_{road}, observed modal share → parameters

Step 2: GC_{road}, constrained modal share, parameters → $GC_{rail,bottleneck}$

We thus obtain that the average generalised cost of rail transport when saturation rises is given by Equation 2.

$$GC_{rail,bottleneck} = \frac{1}{\lambda} \cdot \left\{ \ln\left[\frac{\exp(\lambda \cdot GC_{road})}{p_{rail,bottleneck}} - \exp(\lambda \cdot GC_{road}) \right] \right\}$$

Equation 2

Logit of step 1 is applied until moment 1, when quantities are known and average GC of the two modes can be estimated. Thanks to this step, using many years as observations, logit parameters can be calculated. Alternatively, if literature parameters are available, the first step can be used instead to estimate the GC_{rail}, which is often harder to estimate than as for road traffic.

Logit of step 2 is applied yearly, since the reaching of capacity ($t = t(Q_{rail,max})$; where t is time and Q is the traffic volume), to estimate which is the average extra-cost that, summed to the unconstrained average generalised cost, would generate the same modal share allowed that year by the constraint. The extra-cost is the difference between $GC_{rail,bottleneck}$ and GC_{rail}. We could say that this is the extra-cost capable of shifting the former marginal user, i.e. the user with the less costly alternatives or the user with the lower utility from passing through the network section under consideration.[1]

---

[1] It is worth mentioning a paradox that might rise in the application of the logit. When we calculate the costs of the capacity constraint on the less used mode ("rail", accounting for, say, 30%) we see that the average generalised cost of the other mode ("road", accounting for, say, 70%) is already lower (otherwise "road" would be less used). But, when we force the modal shift to simulate the effect of saturation, the difference in generalised costs (GC_{road} - GC_{rail})

Figure 3 - Generalised costs under saturation conditions (A) and modal shares (B)

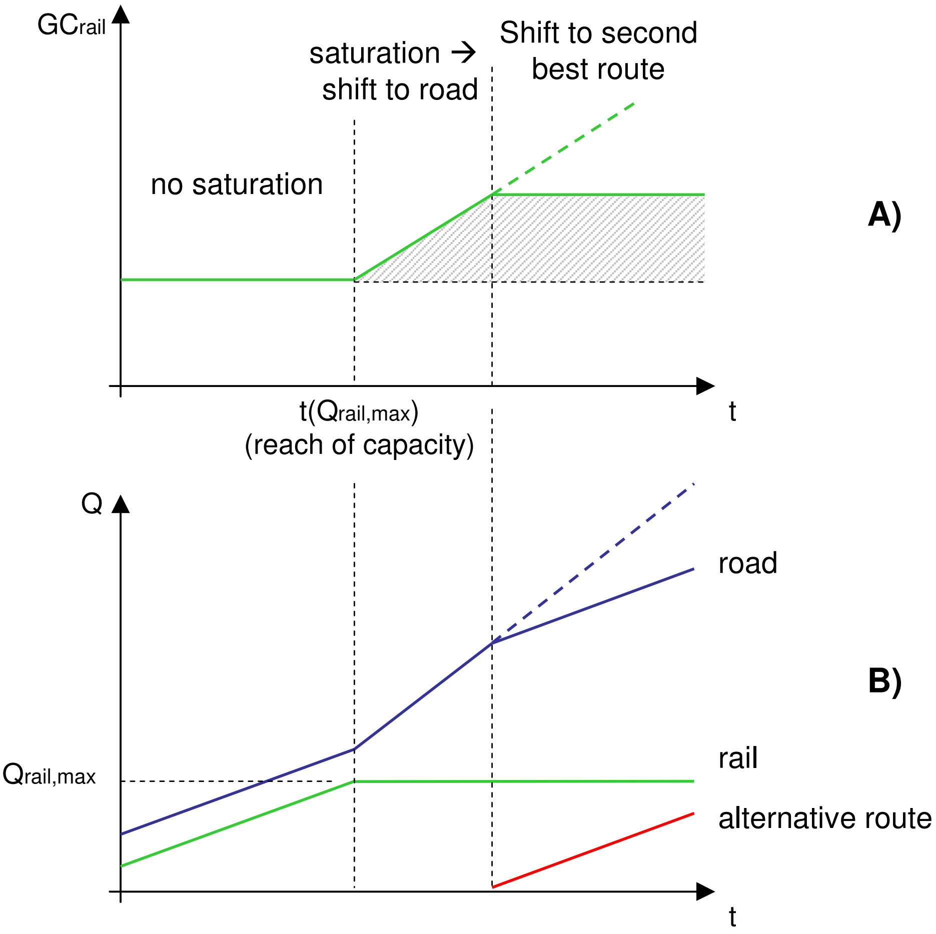


Figure 3 summarises the simulated effect. The imposing of a capacity constraint to the mode "rail" ($Q_{rail,max}$) translates into an increase in the GC of rail[2] (graph A). Consequently, rail traffic stops and levels to the maximum capacity (graph B). This situation is however not infinite (even assuming infinite capacity of "road"), because at a certain point a second best "alternative route" (for example another longer rail path) could become favourable to shifters (graph B). The effect is that the generalised cost increase is capped (shaded are in graph A).

$$GC_{rail,bottleneck,i} = \min\begin{cases} \dfrac{1}{\lambda}\cdot\left\{\ln\left[\dfrac{\exp(\lambda\cdot GC_{road})}{p_{rail,bottleneck,i}} - \exp(\lambda\cdot GC_{road})\right]\right\} \\ GC_{rail\_alternative\_path} \end{cases}$$

Equation 3

However, as it is not known who will actually shift and from where is it coming, it would be impossible to estimate their new costs without further assumptions. What is known is that, coming up to the saturation

---

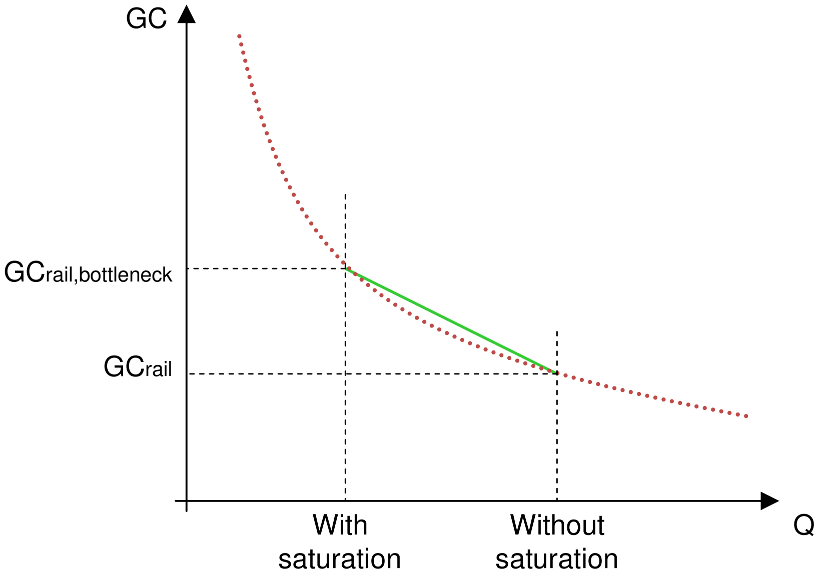
would be positive, resulting in a benefit of the constraint. This paradoxical result is because the existing users of the "rail" already individually perceive lower generalised costs with respect to average ones and consequently use the mode that has the lower cost for them, even if not lower on average. This paradox is however not a problem if the logit is correctly calibrated, typically introducing a modal constant or assuming (as a parameter) that part of the traffic using the dominant mode is captive, both determining generalised costs coherent with previous calculations.

[2] We will discuss later what this increase actually is (a true price or a rent), according to market conditions.

of one mode (say, the rail), the users that have the less expensive alternative or lower utility will be those shifting/disappearing first[3], starting from the former marginal user.

This is summarised in the so called "rule of half", which is the way commonly used in CBAs to overcome the ignorance about the shape of the demand function (Abelson and Hensher, 2001; Kidokoro, 2004; the World Bank, 2005). It assumes that all users are linearly distributed between the former marginal user in the initial situation and the new marginal user in the final situation.

Figure 4 - The "rule of half "



In other words, we do not know for each user its origin and destination, nor the characteristics of the upstream and downstream network. We just know that a certain amount of users would pass with the bottleneck while more would pass if the bottleneck were removed. In both situations, we can measure the costs of the two marginal users.

In conclusion, the logit tells us what is the average generalised cost of transport in both situations and the "rule of half" subsumes the distribution of the non-marginal users.[4]

A similar approach is used by Jorge and de Rus (2004) to evaluate capacity expansion projects in airports.

1.3 The scarcity cost of a bottleneck and the benefit from its removal: user surplus

Our assumptions allows to calculate yearly which is the extra-cost associated to a congested network segment that, if imposed to the users, would cause a modal split compatible with bottleneck capacity constraint.

The difference between the normal generalised cost (GC_{rail}) and the one that simulates the effect of the capacity constraint ($GC_{rail,bottleneck}$), is the extra cost for user, associated to the reaching of capacity. This extra-cost should be applied as a whole to the traffic that remains on the saturated segment, and through the above introduced "rule of half" to the traffic that shifts to the other mode (Figure 5). The "rule of half" subsumes that we don't know how much the shift will cost them, but we know that it will cost something in

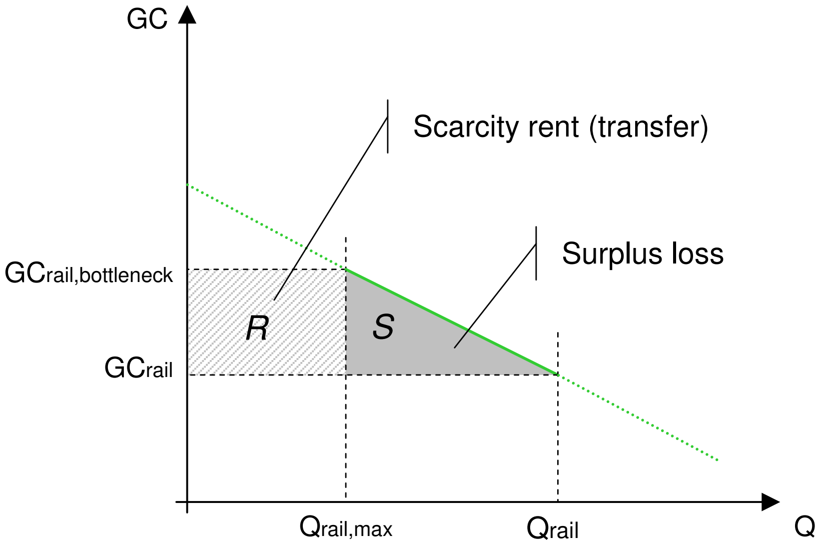
---

[3] We assume here that no grandfathers' right exist: the track operator will price differently in case of saturation and users will modify their behaviour consequently, irrespectively of their previous use of the infrastructure. We will discuss further this hypothesis in Section 5.

[4] Clearly, both hypotheses are strict. However, the case of absence of any reliable transport model to describe the costs of single users is a common situation in practice and this ignorance would not allow to make a correct CBA.

between the difference between the GC of the rail before and after reaching the capacity (otherwise they would not have shifted).

Figure 5 - Shadow cost of bottleneck (at year *i*)



The cost of constraint for the users at year *i* is then:

$$\Delta S_{users,i} = \boldsymbol{R} + \boldsymbol{S} = Q_{rail,max} \cdot (GC_{rail} - GC_{rail,bottleneck,i}) + ½\,(Q_{rail} - Q_{rail,max}) \cdot (GC_{rail} - GC_{rail,bottleneck,i})$$

This extra-cost is however for the users and does not match with the social cost of the bottleneck. We can divide it into two components: a rectangular part (area “R”, the shaded one in Figure 5) is the cost that remaining users “pay” to stay on the saturated line (as a real extra-toll or as a scarcity rent). This, however, is a transfer between the users (more precisely: goods or passengers transported on the line) and another subject (the train operating company, the network manager or the government, according to regulation. See Section 5).

The second component (area “S”, the grey triangle in Figure 5) is instead the net surplus loss associated to the users that renounce to travel on the segment (shifting to another path or giving up on travelling).

The social surplus loss of the bottleneck at year *i* is then:

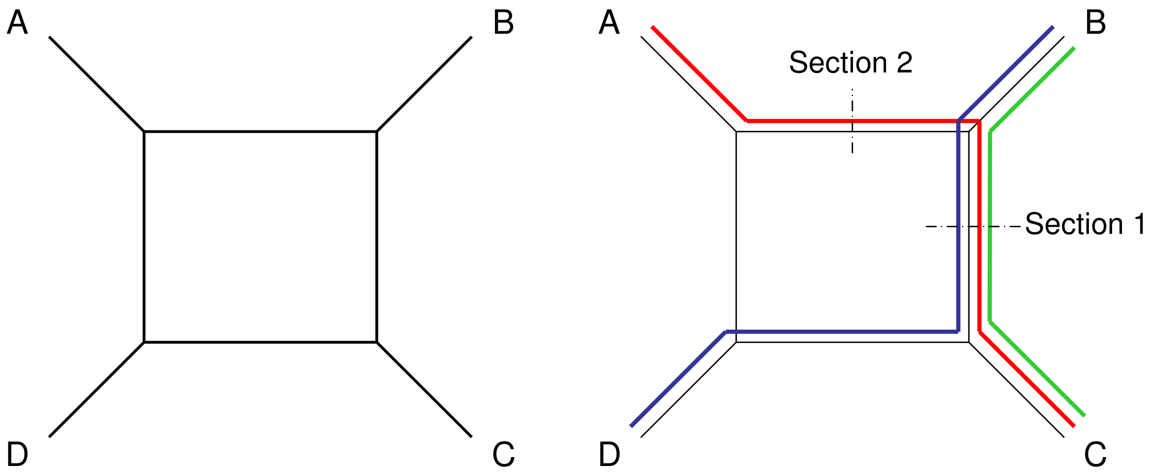
$$\Delta S_i = \boldsymbol{S} = ½\,(Q_{rail} - Q_{rail,max}) \cdot (GC_{rail} - GC_{rail,bottleneck,i})$$

This will be one of the benefits when performing the CBA (see Section 6)

4. Extended case: networks

The same methodology described above for a single segment, can be extended to a network. In this case every segment is to be treated singularly as above, but the way the different traffic components are allocated to the network must follow an optimisation criterion to be decided and that can introduce some complexity in the calculation.

Figure 6 - Example of network structure (left) and example of traffic components' paths (right)



In the left part of example Figure 6 we see a network made of four segments and connecting four origins/destinations. The OD matrix is made of 16 cells and possible paths are two per OD pair. In the right part of Figure 6 an example with just three active relationships (A-C, B-C, B-D) is depicted, each one following a path involving some segments of the network. In particular, one segment is interested by all three OD pairs and two segments are interested by two OD pairs each.

Three different situations may rise.

The first is when one single section (say Section 1 of Figure 6) is saturated by the three traffic components. This case is however trivial, as it must be treated exactly as described in previous Section 3, with the sole difference that the same extra-cost (to be calculated) is applied to three traffic components (via three different logit models) under one capacity constraint.

The second and third case are when two sections (say Section 1 and Section 2 of Figure 6) reach the capacity during the analysed period. In both cases the shadow cost of bottleneck must be calculated as follows:

a) Each OD pair has a fixed average generalised cost $GC_{o,d}$, to which one must add the extra-costs (K_x) to be calculated, one per each section of the network (equal to 0 if no saturation rise), every year.

$$GC_{o,d,bottleneck} = GC_{o,d} + K_1 + K_2 + \dots$$

b) setting all $K_x = 0$ determines the rise of saturation for the two segments, when the logits (one per OD pair) are performed.

c) Starting from the most saturated section[5] (say Section 1), the corresponding extra-cost K_1 is calculated in order to have 100% use of capacity ($Q_{o,d} = Q_{max,o,d}$).

In the second case, this extra cost of Section 1 causes a reduction of traffic such as also the other Section 2 falls below saturation. It must be noticed that the model "automatically" determine the new mix of the traffic components (A-C, B-C, B-D in the example), according to their overall generalised cost, once the cost of constraint is included. So it might also happen that all the shifted traffic come from one single OD pair or any other distribution depending on relative GCs.

In the third case the extra-cost K_1 is not sufficient to solve all upstream and downstream saturation problems, including that of Section 2. In this case a second extra-cost K_2 must be calculated and applied to the involved traffic components (only A-C in this example). As the process described is static, this will however lead to a further decrease of flows on Section 1 (now below 100%). Stopping here would clearly

---

[5] This algorithm is the simplest. Other optimisation algorithms may be applied.

provide an overestimation of the cost of the bottlenecks. So, the process can be repeated iteratively (step 3 will decrease the K_1 previously calculated), until equilibrium is reached with all $Q_{o,d} = Q_{max,o,d}$.

5. The effect of market conditions on users surplus

As previously mentioned in Paragraph 1.3, the distribution of the extra-cost calculated above depends on the market conditions. In particular, it depends on the presence of grandfathers' rights in the access to the saturated network and on the way access is priced to the final users. Naming the two components of users' surplus as in Figure 5 (***R*** and ***S***), we distinguish into 4 cases, summarised in Table 1.

Table 1 - Distributional effect of regulation (sign + is a benefit)

case	*1.1*	*1.2*	*2*	*3*
Cost of scarcity for final users	-R		-R	-R
Extra revenues for TOCs	+R			
Extra revenues for the IM			+R	
Extra revenues for the Regulator				+R
Net surplus variation	-S	worse than -S	-S	-S
ΔS	-S	worse than -S	-S	-S

1) There is no access regulation and the Infrastructure Manager ("IM") does not exploit the cost of scarcity and thus loses the willingness to pay of users, pricing slots at the same price of unsaturated conditions (this happens for example when applying the "grandfathers' rights" principle). In this case it might happen that:
    1.1) Train Operating Companies ("TOCs") are able to discriminate their users rising the price and enjoying of monopoly conditions. In this case final users (the passengers or the transported goods) pay and the TOCs benefit of the same amount as a monopoly rent.
    1.2) TOCs do not discriminate users with higher WTP and go on with their "historical" customers applying "historical" prices up to the reach of saturation, whatever is the WTP of lost demand. From the social point of view this is the worst case, but a quantification of surplus loss is not possible as the WTP of included and excluded demand is not known.
2) There is no access regulation, but the IM is able to exploit the cost of scarcity, by pricing differently the scarce slots. The TOCs and their customers will pay and the IM will get the extra revenues, generating monopoly extra-profits.
3) The Regulator applies scarcity prices for the access to the slots, in order to reflect the calculated cost of scarcity: in this case the demand with higher WTP (through the TOCs) will pay for the extra charge. The remaining demand, with lower WTP from passing in the saturated section, will give up or will find another route. The regulator will re-invest the revenues (not necessarily to solve this bottleneck, that might not be the most socially viable option). This situation is not common, but the most adherent to the effect accounted theoretically by CBA. It is worth noticing that the regulator revenues are not a net benefit, because another component of the market (the users) are paying for it.

It is clear that in all cases what changes is the distribution of the "R" component among the actors of the society according to regulatory conditions. However, in all cases there is a constant surplus loss due to scarcity, that is the component "S" and that represents the surplus loss for the excluded demand. Only in case 1.2 the surplus loss is higher than S, because excluded customers are not necessarily the ones with lower WTP (as happens in other cases). This situation is however very unrealistic, as the TOCs will be probably able to recognise that some of the customers that they do not want to serve are willing to pay more than the "historical" ones.

In conclusion, these different conditions affect the subject that will "pay" the extra-cost of bottleneck. However, this – if the hypothesis of standard CBA theory holds – has no effect on the social surplus, with the relevant exception of the worst case 1.2.

6. A CBA of bottleneck removal projects

In the former sections we have suggested a concise way to evaluate the core benefits of bottlenecks removal in (rail) transport, that is the removal of user surplus loss due to scarcity problems. However, this is not the only aspect related to scarcity problems; moreover, as we said in section 2, a parallel congestion problem exists. In Table 2 we which are the costs and benefits typically related to bottlenecks removal projects, in addition to the surplus loss calculated before (point 5 of Table 2).

Table 2 - General list of costs and benefits to be considered in the appraisal of a rail bottleneck removal

Costs	Benefits
1. Investment costs; 2. New infrastructure operating and maintenance costs; 3. Residual value of the investment. 4. Possible impacts during construction (with respect to existing rail services, other infrastructures and the environment)	5. Removal of scarcity problems: - Surplus gain for users that would be otherwise excluded from the saturated rail infrastructure (as calculated before); - Saved net external costs due to users that would otherwise shift to other modes or longer paths; - Possible wider economic effects. 6. Removal of congestion problems: - improvement in rail service performances (regularity, speed, etc.);

Here some comments about the other benefits to be added. The cost side (points 1 to 4 in Table 2) is instead quite typical and does not need further comments.

Scarcity problems may cause a shift of some users to more impactful modes or to longer paths after the saturation of the existing infrastructure. This will in turn generate external costs (more accidents, congestion, air and noise pollution, climate change, etc.), if not correctly internalised, to the society. The amount is the difference between gross external costs generated by all the shifting users, before and after the project, and new transport specific taxes (that is, fuel duty but not generic VAT[6]) paid by them. It is worth noticing that, while users' surplus variation is calculated for the bottleneck section of the network only (the up- and downstream effects are taken into account with the "rule of half"), the external costs minus the taxes must refer to the entire path followed by the shifting user, that is usually much longer. Reference values for unit marginal external costs in the EU can be found in CE Delft (2008), while transport specific taxes can be estimated referring to Hyelén et al. (2013) for tax rates and the British *Transport Appraisal Guidance* (DfT, 2013: Unit 3.5.6) for fuel consumption formulas.

Another voice of benefits caused by scarcity problems that might play a role in bottleneck removal projects is that of wider economic effects (point 8 in Table 2). Wider economic effects represent all those benefits that are not captured by direct benefits to users in a well constructed standard CBA, after allowing for external costs (Vickerman, 2007): literature includes agglomeration economies, increased competition, productivity of firms and network effects. These effects – that can be considered to be negligible in many transport projects – might potentially be significant in bottleneck removal projects: the sharp increase in

---

[6] Actually, VAT applied on fuel duty represent a transport specific tax and should be considered.

perceived generalised costs of transport due to the reaching of capacity – that, as we have seen, partially represents a transfer within the society – might trigger effects which are not transfer.

Unfortunately, the evaluation of wider economic effects usually required complex economic model. An alternative way is the one suggested in the British *Transport Appraisal Guidelines* (DfT 2013: Unit 3.5.14), but it still needs a lot of detailed and geographically disaggregated data. A parametric estimation of what should be the WEE capable of justifying an otherwise negative CBA of bottleneck removal is thus a more manageable way to give policy maker an answer.

We finally have to consider possible congestion problems, particularly if the chosen maximum rail capacity value is not conservative. Regardless what criterion has been used to determine this value, the performance of rail services will gradually worsen when the number of trains will get closer to this value, in terms of less regularity of the service (more delays) and possible speed reductions. The increase of maximum capacity will thus provide also benefits to the users of those services, in particular passengers, and these must be included in the final NPV. If the capacity value is defined with respect to an average delay value, this can be useful to make an early appraisal.

7. Concluding remarks

In this working paper we proposed a simplified approach to estimate the effects of a capacity constraints for one or more segments of a rail network, and assess its removal through a CBA. Using binomial logit models, we can assess the change in the generalised costs of rail transport under saturated conditions and thus evaluate the scarcity costs for the remaining users (which represent a transfer within the society) and the surplus loss of users excluded from the saturated rail infrastructure (which is instead a net social cost of the bottleneck). The paper is supplemented with indications about how to evaluate the other possible benefits of bottleneck removals, that is improvement in rail service performances, saved net external costs due to users that would otherwise shift to other modes or longer paths and possible wider economic effects.

8. Bibliographical references

Abelson, P. and P. Hensher. 2001. "Induced travel and user benefits: clarifying definitions and measurement for urban road infrastructure," in Hensher D. and K. Button (eds.): *Handbook of transport systems traffic control*, Handbooks in Transport 3. Elsevier, Pergamon;

Cascetta, E. (2001). *Transportation systems engineering: theory and methods,* Vol. 49, Springer.

CE Delft (2008). Handbook on estimation of external costs in the transport sector. *Internalisation Measures and Policies for All external Cost of Transport (IMPACT). Version*, *1*.

DfT (2013), *Transport Analysis Guidance*, Department for Transport, London (UK). Website: http://www.dft.gov.uk/webtag

EC-EIB (2005), *RailPAG: Railway Project Appraisal Guidance*, European Commission and European Investment Bank.

Eliasson, J., & Borjesson, M. (2012). On timetable assumptions in railway investment appraisal. In *Transportation Research Board 91st Annual Meeting*(No. 12-2124).

Jorge, J., D., & de Rus, G. (2004), "Cost–benefit analysis of investments in airport infrastructure: a practical approach", *Journal of Air Transport Management*, 10, 311–326.

Kidokoro Y. (2004). Cost benefit analysis for transport networks. Theory and application. *Journal of transport Economics and Policy*, vol. 38, part 2, 275-307.

Nash, C., & Sansom, T. (1999). *Calculating transport congestion and scarcity costs*. Final Report of the Expert Advisers to the High Level Group on Infrastructure Charging. Commission of European Communities, Brussels.

Network Rail (2009), *New Lines Program: Strategic Business Case*, Planning and Regulation, Route Planning, London (UK).

Nilsson, J. E. (2012). *Congestion and scarcity in scheduled transport modes.* CTS Working paper No. 2012: 25. CTS-Centre for Transport Studies Stockholm (KTH and VTI).

Hylén, B., Kauppila, J., & Chong, E. (2013), *Road Haulage Charges and Taxes: Summary analysis and data tables 1998-2012*, Discussion Paper 2013 – 08, OECD/ITF.SRA (2003), *Sensitive Lorry Miles: Results of Analysis*, Strategic Rail Authority, UK.

The World Bank (2005), *Transport Note No. TRN-11: treatment of induced traffic*, Notes on the Economic Evaluation of Transport Projects, Transport Notes, Transport Economics, Policy and Poverty Thematic Group, Washington, DC (USA).

Vickerman, R. (2007), *Recent Evolution of Research into the Wider Economic Benefits of Transport Infastructure Investments*, Discussion Paper No. 2007-9, ITF-OECD Joint Research Centre.