On timetable assumptions in railway investment appraisal

Jonas Eliasson and Maria Börjesson

Published in Transport Policy, vol. 36, p. 118-126, 2014

Abstract
The benefits captured in an appraisal of a railway investment are determined by what timetables the analyst assumes in the scenarios with and without the investment. Without an explicit, objective and verifiable principle for which timetables to assume, the appraisal outcome is virtually arbitrary. This means that appraisals of railway investments cannot be compared to each other, and opens the door for strategic behaviour by stakeholders conducting seemingly objective cost-benefit analysis. We explain and illustrate the nature and extent of the problem, discuss possible timetable construction principles, and show that current practice is likely to exaggerate appraisal benefits.

Acknowledgments
Thanks to Lena Wieweg for fruitful discussions, and to Chris Nash and Jan-Eric Nilsson for useful comments on an earlier draft of the paper. Lennart Lennefors and Magnus Wahlborg at the Transport Administration helped explaining the capacity relationships used by the Transport Administration.

Date: 18 Aug 2014

Keywords: Cost-benefit analysis, appraisal, railway investments, timetables.
1. Introduction

Railway investments are substantial financial commitments. In most countries, developing and maintaining the railroad network is a public responsibility. As with all public transport policy decisions, benefits can be weighed against costs and the cost-efficiency of different investments can then be compared. This makes well-developed and well-structured cost-benefit analysis (CBA) methods indispensable. The increasing interest for investments in high-speed railways makes this even more important, considering the huge costs at stake.

The outcome of a public railway investment appraisal depends on the assumed timetables with and without the investment. Yet principles for timetable design in appraisal are not explicitly discussed in the literature, neither in national appraisal guidelines nor in the research literature. The purpose of this paper is to explain and illustrate the crucial importance of timetable assumptions in public railway investment appraisal, why explicit principles for appraisal timetable construction are necessary, discuss the potential problems that the lack of such principles creates, discuss different such principles, and explain why current practice will tend to exaggerate investment benefits.

The key insight is that without an explicit principle governing the choice of appraisal timetables both with and without an investment, the social benefit of a railway investment is not defined. It is meaningless to speak of the benefits of a railway investment as such – it is the use of the investment, that is, the timetable, which decides what benefits a railway investment will generate.

Railway investments are different from most other transport investments in that its use is often planned and decided by society to a large extent, through capacity allocation, regulations, track access charges, subsidies, public transport provision etc., and by commercial operators, who strive to maximize profits, often with some monopolistic power. This stands in contrast to most other transport investments, such as roads or bicycle paths, where the use of the physical infrastructure is mainly decided by individual decisions made by large numbers of travellers. Which timetable principle is most appropriate to use in an appraisal will depend on the institutional setting – which stakeholders decide the use of the railway in question and their incentives, goals and constraints, and how potential conflicts between stakeholders are resolved. Hence, this paper cannot prescribe what principle should govern the construction of appraisal timetables. The purpose is merely to describe the problem, illustrate its potential consequences, and discuss some principles which may be appropriate. There is a multitude of possible institutional settings, and in this paper we only illustrate a few examples. The insight is quite general, however, and occurs in any institutional setting. However, it is likely to be especially problematic when the responsibilities for financing investments, capacity allocation and running the trains are separated between different stakeholders.

The importance of timetable assumptions becomes particularly evident when appraising capacity improvements, since an increase in capacity can be used in any combination of increased frequencies, decreased travel times, and reduced delays. Depending on how the new capacity is used, an appraisal will give different results. The appraisal outcome will also depend on what is assumed about the use of the existing capacity in the do-nothing case. It is not only the social benefits that depend on timetable assumptions; forecasts of revenues and passenger volumes obviously also do so.
This means that the analyst responsible for carrying out an appraisal can influence how large benefits an investment will generate in the analysis by choosing more or less socially efficient timetables\(^1\). This influence can be conscious or subconscious. It is usually difficult or impossible to gauge the social efficiency of a timetable for outside observers and decision-makers, and sometimes even for the analyst. Clearly, this means that CBA results may become misleading and to some extent even arbitrary. Even worse, if the analyst represents a stakeholder with an interest in discrediting or promoting certain investments, there is an opportunity for strategic behaviour, with very little possibility for outsiders to discover this.

Appraisal timetables for long-term strategic investment planning are typically simplified down to its essentials. The timetables are typically defined only by frequencies and travel times of different train types (e.g. regional, long distance and freight trains), and we will follow this convention in the examples in this paper. However, the level of detail of the timetables does not influence the main insight that the social benefit of an investment depends on the assumptions regarding them.

The intended audience for this paper is not primarily railway planners but people working with appraisal methodology. Appraisal methodology covers all sorts of inputs, valuations and processes, from future fuel prices to the value of a statistical life. To ensure that CBAs for different investments can be compared and ranked consistently, there are often national guidelines for appraisal methods, inputs and parameters. In contrast, principles for timetable construction are surprisingly neglected. The ultimate objective of this paper is to start a discussion about such principles. Again, which timetable principle is appropriate depends on the prevailing institutional setting – the incentives of the involved stakeholders and the process for resolving capacity conflicts. The principle needs to capture the essential interactions between demand, supply, capacity constraints and the capacity allocation process. Once such a principle is established, it can be used to calculate timetables both in the do-nothing and in the investment scenario.

In Section 2, a simple model is formulated, which captures the main features of applied transport CBA. The model is a simplified version of current Swedish appraisal practice and guidelines. Section 3 shows, by way of examples, why the social benefits of a railway investment cannot be defined without a principle for choosing timetables before and after the investment, and that this opens the door for both honest confusion and strategic behaviour by the analyst. Section 4 discusses advantages and disadvantages of different timetable principles. Section 5 discusses the (implicit) principles for timetable construction in current appraisal practice, and what consequences these principles are likely to have. Section 6 concludes.

1.1 Literature and practice

Despite the obvious importance of timetables in appraisal, we have failed to find any mentioning in appraisal guidelines of how analysts should construct timetables. For example, the extensive guideline RAILPAG (Railway Project Appraisal Guidelines), intended to be "a common framework for the appraisal of railway projects across the EU" does not mention how appraisal timetables should be designed, what stakeholder should have the responsibility for it, or how appraisal results might be affected by it. Neither of the national appraisal guidelines of

\(^1\) A socially efficient timetable is the timetable that maximizes the sum of all social benefits and costs, including external costs and benefits. This includes e.g. passenger costs and benefits, producer costs and revenues, emissions and accidents.
Sweden, Norway, UK or the Netherlands mentions the topic, despite having detailed formal appraisal guidelines covering virtually all other parameters and dimensions.

Even if there seems to be an understanding that the timetables determines investment benefits among many practitioners and in the grey literature, the practice in most railway appraisals is to treat the timetable as an exogenous variable outside the analyst’s control or responsibility. This is both strange and problematic, since comparability of appraisals of different investments is essential. For this reason appraisal guidelines devote substantial efforts to forecast and harmonize assumptions regarding virtually all other variables entering the CBA, from future oil prices and population growth to transit fares and vehicle operating costs.

It is true that the future timetable - the timetable which the appraised project will enable – often attracts interest, even if there is seldom any explicit principle guiding its construction. However, much less thought is usually put into constructing the timetable in the do-nothing scenario, despite the fact that it is well known that the definition of the do-nothing scenario is always crucial in any CBA exercise. That scenario should be the best possible alternative if the project is not carried out. Although this principle is often difficult to apply in practice, the idea can be operationalized by evaluating many alternatives and rank them mutually. Van Wee (2007) discusses the importance of do-nothing alternative in the context of railway CBA further.

There is no research literature that explicitly deals with the specific problem of how the social benefit of a rail investment depends on the assumed timetables in the scenarios with and without the investment. But there are several branches of literature dealing with related problems.

First, there is a body of research literature on demand modelling, showing that behaviour changes in response to levels of rail service and timetables (e.g. Hensher (1997) and Hensher and Ton (2002)), but in this literature the timetables arrive by assumption. This literature does not consider the problem of the present paper – how the choices of timetables influences appraisal.

Second, there is a developing literature concerning different aspects of railway operations optimization, such as crew and vehicle scheduling and train scheduling and routing. This literature typically uses engineering objectives such as minimizing total operations cost (see Caprara, Kroon, Monaci, Peeters & Toth (2007) for an overview). Social welfare or any other economic objective function, however, is typically not considered in this literature. Demand is usually handled simplistically, and if passenger welfare is taken into account, this is done in a crude or heuristic fashion such as maximizing the number of passengers able to reach their destination without an interchange. A rare example of an economic objective function is Brännlund, Lindberg, Nõu and Nilsson (1998), who present an algorithm that schedules a set of trains to obtain a profit-maximizing timetable without violating track capacity constraints.

Third, there is some literature that at least implicitly indicates the specific problem of the present paper. Adler, Pels and Nash (2010) study how competition between airlines and high-speed rail may affect service levels, and how this affects the benefits of high-speed rail investments. There is an insight of the problem that the assumed timetables affect the utility of rail investments although no explicit discussion of or solution to it, such as an explicit principle for timetable construction in appraisal, is suggested. Bristow, Preston and Nash (1998), Nash and Preston (1991) and Nash (1992) all touch upon related issues when discussing and analysing
appraisal principles in the partially deregulated British railway industry. However, there is no explicit formulation of the problem or suggestions how it can be solved.

Fourth, ex-post studies of rail investments confirm the importance of timetable assumptions, concluding that they are essential for demand forecasts and thus for social benefits, and an important source of forecasting errors. For example, Tegnér (2001) concludes that the optimistic timetable (high frequencies, low fares, short travel times) used in the demand forecast for the Ostkustbanan railway line was a major reason for a fourfold overprediction of increase in demand. SIKA (2005) found that the actual number of travellers on three studied Swedish railway lines was overpredicted with between 26 per cent and 72 per cent, quoting wrong timetable assumptions as one of the reasons for the demand shortfall. Olsson (2006) and Mackie and Preston (1998) also point out that the accuracy of the calculated benefit of rail investments relies on the assumptions of the realized timetable.

2. The model

In this section we set up a model framework for cost-benefit analysis of rail capacity extensions. In section 3 we will apply this model to some numerical examples illustrating why the benefit of a rail investment cannot be defined without an explicit principle for timetable construction.

For clarity, the model only includes the essentials of a CBA: a demand function, operator production costs, and a capacity relationship describing how train running times depend on the scheduled number of trains and the infrastructure. The nonlinear demand and linear supply functions, and their parameters, are taken from Swedish rail guidelines, so they are perfectly realistic albeit simplified to only one class of travellers. The functional forms do, however, not matter for the insights of the analysis. Capacity relationships can be modelled in different ways, depending on the context and the need for details. In one extreme, real, detailed timetables are constructed, which are then analysed graphically or through simulation. In the other extreme, aggregate, analytical relationships based on the capacity relationships are used. The aggregate approach is usually used for strategic planning and appraisal. We apply the capacity relationship used by the Swedish Transport Administration for strategic planning and appraisal, which is similar in structure to corresponding capacity relationships used in other countries. This relationship is defined in section 2.1. The timetables in the model are simplified down to the essentials: frequencies and travel times in a homogeneous time period. Assuming schedule-based timetables, rather than frequency-based, would just add complexity to the exposition without adding further insights.

2.1 The capacity relationship

Consider a line segment between two stations of length $L$. Let $n_i$ be the total number of trains of train type $i$ in both directions traversing this line segment during the analysed time period $S$ (for example a rush period or a day). Symmetry is assumed,

---

2 The same trade-off between detail/realism and ease-of-use/analytical tractability exists for road capacity relationships, where there are also both dynamic and static volume-delay relationships.

3 This relationship is in fact not very realistic: it tends to underestimate travel times at high capacity utilization. Nevertheless, it is sufficient for illustrating the points we wish to make. Besides, it is of practical importance, since the relationship is used in Swedish practice.
implying that the number of trains in each direction is $n_i/2$. All trains of type $i$ have travel time $t_i$ and headway $2S/n_i$.

The consumed capacity $C$ is the fraction of the analysed time period $S$ when the line segment is occupied by trains. Each line segment is divided into homogenous track segments $k$, which are the smallest building blocks of the network. $C$ is calculated for each track segment on the line segment, and the highest $C$ is applied to the whole line segment. For single tracks, $C$ is defined as

$$C = \max_k \left[ \sum_i (n_i T_{ik} + \rho n_i M_i) / S \right],$$  

(1)

where $M_{ik}$ is a time addition for meetings, depending on train type $i$, $T_{ik}$ is the minimal travel time of track segment $k$ for train type $i$ and $\rho$ is the probability that two trains meet. The consumed capacity is thus the fraction of the time period $S$ during which a train occupy the segment, including the waiting time for meetings. The scheduled travel time for trains of type $i$ on the line segment, $t_i$, is then computed from the minimal travel time of the line segment, $T_i$, track length $L$ and $C$:

$$t_i = T_i + \max\{0.2C - 0.06L, 0\},$$

(2)

where $L$ is given in kilometres and $T_i$ in minutes. This particular functional form and parameters are taken from the Swedish guideline for strategic planning and appraisal (Swedish Rail Administration, 2005), which is based on practical experience. They are obviously simplistic and not applicable for more detailed capacity calculations, but as noted above, the specific functional forms do not matter for the fundamental insights of the analysis.

### 2.2 Demand and consumer surplus

The demand model captures the essential features of a transport CBA but is as homogeneous as possible: it considers only one class of travellers and one ticket price. For simplicity, the fare $p_i$ is assumed to be constant and exogenous.

4 This is consistent with appraisal practice, although often not realistic. With a constant elasticity demand function the optimal price, $p^* = (\gamma_2 t + \gamma_3 L) \varepsilon / (1+\varepsilon)$, is independent of both the number of trains $n$ (as long as $n$ is independent of $t$) and initial demand $D_0$. With for example a linear demand function, this is no longer true.

5 The crowding cost is not included because it is implicitly assumed in the operations cost relationship that the trains become longer as the number of passengers increases so that crowding is avoided. This is to avoid having endogenous generalized travel costs, which would complicate things without adding more insight in the examples.

The generalized travel cost is defined as $c_i = p_i + \alpha t_i + \beta S/n_i$, where $S/n_i$ is the average waiting time, and $\alpha$ and $\beta$ are monetary values of travel time and waiting time. Travel demand between the stations is

$$D_i = D_i(c_i).$$

In the numerical examples we will assume a demand function with constant elasticity $\varepsilon$ such that $D(c) = D_0 \varepsilon (c_0/c)^\varepsilon - 1$. The consumer surplus (CS) of a change in generalized costs from $\{c_i^0\}$ to $\{c_i^1\}$ is by definition:

$$CS = \sum_i \int_{c_i^0}^{c_i^1} D(c) dc \approx \sum_i \frac{1}{2} \left[ D(c_i^0) + D(c_i^1) \right] (c_i^1 - c_i^0).$$

(3)
In the numerical examples, we will use the rule of a half approximation, which is valid for small changes \((c_i^0 - c_i^1)\).

### 2.3 Operator costs and revenues

The producer surplus (PS) is defined as fare revenues minus operations costs for the operator. Operations costs increase with the number of trains, the number of passengers, travel distance and travel time. We use a cost function used in Swedish appraisal, where operations costs are assumed to be linear in train operating time \((n_i t_i)\), passenger time \((D_i t_i)\) and passenger distance \((D_i L)\). The producer surplus of a change of travel times and the number of trains from situation 0 to situation 1 is then:

\[
PS = \sum_i (D_i^1 - D_i^0) p_i + \gamma_1 (n_i^0 t_i^0 - n_i^1 t_i^1) \\
+ \gamma_2 (D_i^0 t_i^0 - D_i^1 t_i^1) + \gamma_3 (D_i^0 - D_i^1)L
\]

where \(\gamma_1-\gamma_3\) are parameters. Total social benefits are defined as \(CS + PS\).

### 3. Numerical examples

In this section, we will illustrate how the outcome of a CBA is almost arbitrary when there is no explicit principle deciding how timetables should be determined (both in the do-nothing and the do-something scenarios). This is quite general; it is true regardless of the institutional setting, i.e. which stakeholders decide the use of the railway and what objective functions they have. Which timetable principle is the most appropriate will depend on this institutional setting. There is obviously a host of possible such settings; here, we will only provide two simple examples to illustrate the problem.

#### 3.1 One train type

Consider a 100 km single track line segment between two stations, divided into 10 track segments of equal lengths, with meeting stations between each track segment. To demonstrate how the benefit of an investment may vary depending on timetable assumptions in the CBA, we will now apply the model set up in the previous section to compute the benefit of a suggested capacity extension on this track line segment. The example is realistic but simplified. Parameter values are adapted from the Swedish appraisal guidelines and given in Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value of in-vehicle time</td>
<td>70 SEK/h</td>
</tr>
<tr>
<td>Value of waiting time</td>
<td>70 SEK/h</td>
</tr>
<tr>
<td>(\gamma_1)</td>
<td>89.1 SEK/min/train</td>
</tr>
<tr>
<td>(\gamma_2)</td>
<td>0.266 SEK/min/pass</td>
</tr>
<tr>
<td>(\gamma_3)</td>
<td>0.102 SEK/km/pass</td>
</tr>
<tr>
<td>Demand elasticity</td>
<td>-0.7</td>
</tr>
</tbody>
</table>

\[
D(c) = D^0 \left(\frac{c}{c^0}\right)^{-0.7}
\]

where \(D^0\) and \(c^0\) is the initial situation.

According to this guideline, the value of waiting time is the same as the value of in-vehicle time for waiting times under 15 minutes, and lower for longer waiting times.

Assume that there is only one train type, with a minimum travel time of \(T = 40\) minutes on the line segment and \(T_k = 4\) minutes on the track segments. Assume that there is initially \(n=4\) trains within a time period \(S=60\) minutes, such that...
we have 4 trains/hour passing between the two stations (2 trains/hour in each direction). Initially they carry $D_0=1000$ passengers/hour. The probability that a train meets another train at any given track segment is

$$\rho = n \cdot \frac{T}{L} = 4 \cdot \frac{40}{100} = 0.27.$$

With $M=5$ minutes, we have from eqn. (1):

$$C = \max_k \frac{n_i T_{ik} + \rho n_i M_i}{S} = \frac{4 \cdot 4 + 0.27 \cdot 4 \cdot 5}{60} = 0.36,$$

and we have from eqn. (2):

$$t_i = T_i + \max \{ (0.2 \cdot C - 0.06) L, 0 \} = 40 + \max \{ (0.2 \cdot 0.36 - 0.06) \cdot 100, 0 \} = 41.11 \text{ minutes}.$$

To improve the service it is suggested that the future frequency should be 7 trains/hour. To avoid increased travel times, it is suggested that the capacity is increased by doubling the number of meeting stations is doubled, making each track segment 5 km long. With this capacity extension, a frequency of 7 trains/hour results in a travel time of 41.4 minutes (using the capacity equations (1) and (2)), i.e. virtually unchanged compared to the current situation. Now, a CBA of this capacity expansion is to be set up. In the investment scenario, a timetable of 7 trains/hour is assumed. It is less obvious what timetable should be in the do-nothing scenario, however. Assuming the current timetable, 4 trains/hour (giving a travel time of 41.4 minutes), results in appraisal A shown in Table 2. The consumer surplus (CS) of this capacity increase is computed by eqn. (3) and the producer surplus (PS) is computed by eqn. (4), with parameters given in Table 1. The total social benefit (CS+PS) becomes slightly more than 7000 SEK/hour.

Table 2. Example 1, appraisal A.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Investment</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trains/hour</td>
<td>4</td>
<td>7</td>
<td>-3</td>
</tr>
<tr>
<td>Travel time (minutes)</td>
<td>41.1</td>
<td>41.4</td>
<td>0</td>
</tr>
<tr>
<td>Passengers/hour</td>
<td>1000</td>
<td>1085</td>
<td>-85</td>
</tr>
<tr>
<td>Generalized cost (minutes)</td>
<td>65</td>
<td>58</td>
<td>7</td>
</tr>
<tr>
<td>Consumer surplus (SEK/h)</td>
<td></td>
<td>7480</td>
<td></td>
</tr>
<tr>
<td>Producer surplus (SEK/h)</td>
<td></td>
<td>-332</td>
<td></td>
</tr>
<tr>
<td>Total social benefits (SEK/h)</td>
<td></td>
<td>7148</td>
<td></td>
</tr>
</tbody>
</table>

Another analyst, however, argues that since the future frequency of 7 trains/hour is planned or needed either way, this is the relevant timetable also in the do-nothing scenario. In this CBA, summarized in Table 3, the total social benefit becomes approximately 28 000 SEK/hour, about four times the social benefits in the first analysis.
Table 3. Example 1, appraisal B.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Investment</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trains/hour</td>
<td>7</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Travel time (minutes)</td>
<td>48.8</td>
<td>41.4</td>
<td>7</td>
</tr>
<tr>
<td>Passengers/hour</td>
<td>985</td>
<td>1085</td>
<td>-100</td>
</tr>
<tr>
<td>Generalized cost (minutes)</td>
<td>67</td>
<td>58</td>
<td>9</td>
</tr>
<tr>
<td>Consumer surplus (SEK/h)</td>
<td></td>
<td></td>
<td>8913</td>
</tr>
<tr>
<td>Producer surplus (SEK/h)</td>
<td></td>
<td></td>
<td>19401</td>
</tr>
<tr>
<td>Total social benefits (SEK/h)</td>
<td></td>
<td></td>
<td>28314</td>
</tr>
</tbody>
</table>

Then, a third analyst argues that the investment should be evaluated keeping the frequency fixed at 4 trains/hour, arguing that this is the “pure” investment benefit. This reduces social benefits by almost a half compared to the first result, to less than 4 000 SEK/hour (appraisal C in Table 4).

The question of which of these analyses is correct cannot be settled without a principle determining how timetables are constructed, and using the same principle both with and without the investment. One possible principle is to assume that frequencies are set to maximize social benefits; another principle is to assume that they are set to maximize operator profits. Several other principles are imaginable.

Table 4 summarises investment benefits under a number of different timetable assumptions. Appraisal A (row 1) compares the current situation with the planned situation. This is a reasonable alternative, although it potentially adds the effects of the investment as such with the effect of the timetable change as such. Appraisal B (row 2) assumes that train frequencies are increased regardless of whether the investment is built. This is a common assumption in appraisal practice: a planned target timetable is assumed, and investments are evaluated assuming the same timetable. Appraisal C evaluates the investment at the current frequency, thus separating the effect of the increased frequency from the effect of the investment, although obviously not using the new capacity to its full extent. Appraisal D and E assume that train frequencies are set to maximize operator profits and social benefits, respectively, in both scenarios.

Table 4. Example 1: Investment benefits assuming different timetables.

<table>
<thead>
<tr>
<th>Principle</th>
<th>Frequencies</th>
<th>CS</th>
<th>PS</th>
<th>CS+PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Current vs. planned frequencies</td>
<td>( n^0=4, n^1=7 )</td>
<td>7480</td>
<td>-332</td>
<td>7148</td>
</tr>
<tr>
<td>B. Planned frequency</td>
<td>( n^0=7, n^1=7 )</td>
<td>8913</td>
<td>19401</td>
<td>28314</td>
</tr>
<tr>
<td>C. Current frequency</td>
<td>( n^0=4, n^1=4 )</td>
<td>1305</td>
<td>2513</td>
<td>3818</td>
</tr>
<tr>
<td>D. Profit-maximizing frequencies</td>
<td>( n^0=3.5, n^1=5.7 )</td>
<td>7992</td>
<td>3691</td>
<td>11682</td>
</tr>
<tr>
<td>E. Welfare-maximizing frequencies</td>
<td>( n^0=4.0, n^1=6.0 )</td>
<td>7429</td>
<td>4421</td>
<td>11851</td>
</tr>
</tbody>
</table>

We will explain in some detail why so different conclusions can be obtained. Figure 1 plots the consumer surplus (CS) (see eqn. (3)), producer surplus (PS) (see eqn. (4)) and total social benefits (CS+PS) as a function of the number of trains in the initial situation without investment. All benefits are compared to the reference situation with \( n=4 \) trains/h.
Figure 1. Example 2: Consumer surplus, producer surplus and total benefits as functions of the number of trains, initial situation without investment

The CS increases with the number of trains up to 5 trains/h, driven by declining waiting time. Above 5 trains/h, the CS declines with increasing number of trains because the travel time increases due to capacity constraints. The PS has a maximum at 3.5 trains/hour; fare revenues increase slower than operating costs after that point. Total benefits PS+CS attain a maximum at 4 trains/h. We can note in passing that the profit-maximizing optimum is different from the social optimum; a commercial operator trying to maximize profits will run fewer trains than what is socially optimal. This is one of the arguments for society providing, regulating or subsidizing public transport.
Figure 2 corresponds to Figure 1, but shows the benefits in the investment scenario. The PS has its maximum at 5.7 trains/hour and the total benefits CS+PS has a maximum at 6.0 trains/hour. The investment increases both CS and PS compared to the reference situation.

Figure 3 shows total benefits (eqn. (1))) with and without the investment, showing Appraisal A and Appraisal B as arrows. The key point is that the net benefit of the investment depends on what is assumed about the number of trains before and after the investment, \( n^0 \) and \( n^1 \). In Appraisal B, the frequency in the do-nothing scenario is far from the social optimum, causing the appraisal to show large benefits. In Appraisal A, the timetable happens to be close to socially optimal in the do-nothing scenario, hence shows much smaller benefits than Appraisal B.

In this simplistic example, this is intuitively obvious. But in a real example, these effects are effectively impossible to see through, especially for a third-party observer. An appraisal might even show negative benefits, if the timetable is further from the optimal in the investment scenario than in the do-nothing scenario. For example, this happens if one assumes 4 trains/hour without the investment and anything above 8 trains/hour with the investment. We are aware of such phenomena happening in practice, confusing the analysts involved.

What timetable principle is appropriate depends on the institutional context. If the institutional setting is such that, say, a public agency strives to maximize social welfare, then the welfare-maximizing frequencies should be applied when evaluating the investment (4 trains/hour in the do-nothing scenario and 6 trains/hour in the investment scenario). If the institutional setting is such that an operator is free to maximize its profits, the number of trains that optimize the producer surplus should
be applied in the appraisal. Appraisals D and E happen to be close to each other and to Appraisal A, but this needs not be the case.

Capacity utilization does not only affect travel times but also delays, since high capacity utilization increases knock-on delay effects. In practice, this is partly compensated for by adding buffer times to the timetable travel times. We have not included delay effects in the example, but analytically we can interpret the travel times in the example as if they include buffer times and average delays (valued by an appropriate multiplier). In the above example we have assumed capacity constraints, but it is easy to show that the analysis is sensitive to timetable assumptions also where there are no capacity constraints, and the investment only reduces travel time. For capacity improvements the analysis is, however, particularly sensitive since the added capacity can allow for different improvements: travel times, frequencies or reliability.

3.2 Two types of trains

Next, assume that there are two operators with different objectives, and infrastructure is funded by the government. This example aims at illustrating how timetables can result from a negotiation or decision process involving several actors, and that operators may have incentives to state a future timetable different from their real intentions. Based on this example we will then discuss what principle for timetable assumptions in appraisal that is appropriate, depending on the institutional setting.

Assume that there are two train types competing for capacity: regional trains and long-distance trains. The two demand functions for the train types are completely separable; we can imagine that the two trains serve separate markets, and only happen to share this particular stretch of track. Regional trains are operated by a public operator that strives to maximize consumer surplus (the regional government covers operating deficits), while long-distance trains are run by a commercial operator trying to maximize profits. The institutional setting is such that the regional train operator first decides the number of trains it will run, and then the long-distance train operator decides the number of trains it will run, given the intentions of the regional train operator. The government funds infrastructure investments, but does not take into account that the timetables assumed for the do-nothing and the investment scenarios affect the outcome of the appraisal. The operators, however, are fully aware of the importance of the timetables used in the appraisal scenarios.

Assume that regional trains are slower, taking 45 minutes to traverse this track if there are no capacity constraints, while long-distance only need 25 minutes. In the current situation, there are 3 regional trains per hour with a total of 1000 passengers, and 3 long-distance trains per hour, also with 1000 passengers. The longest track segment on the 100 km single track line segment between two stations is 15 km. The other parameters are unchanged from the example with a single train type.

Now, the regional train operator wants to increase capacity by increasing the number of meeting stations such that the longest track segment decreases from 15 km to 3 km, and extend the timetable to 6 trains/hour. Given 6 regional trains/h, the long-distance operator would maximize profits by running 5 trains/hour in the new situation. The first row of Table 5 shows the benefits of the investment with the timetables that will be the actual outcome in the do-nothing and the investment scenarios: 87 kSEK/h.

However, since the government does not take into account the impact of the timetables in the appraisal, the operators have both incentive and opportunity to act
strategically when stating their intended timetables, which are inputs to the appraisal. Since the consumer surplus and the producer surplus both increase when the investment is built, the operators support the capacity investment and have incentives to increase the CBA benefits, increasing the probability that the government decides to build the investment.

The operators know which timetable gives the maximal social welfare in the investment scenario (3 regional and 5 long-distance trains), so they can act strategically by stating that they intend to run this number of trains if the investment is built. This yields a benefit of 163 kSEK/h (second row of Table 3), almost twice the benefit of the actual outcome. In the investment scenario, the timetable maximizes the social welfare, but in the do-nothing scenario the train frequency is higher than the optimal.

The operators can also act strategically by stating (truthfully) their intentions to run 6 regional trains/h and 5 long-distance trains/h with the investment. However, they state that they also intend to run this timetable even if the investment is not built. Then the benefit would increase further to 306 kSEK/h.

The resulting benefit obviously depends on the institutional setting. Another possible setting can be that there is one profit maximizing operator that runs both train types. The realized timetables in the do-nothing and investment scenarios would then be the ones that maximize the total PS, yielding an investment benefit of 115 kSEK/h. A third possible setting is that both train types are run by a public authority, maximizing social welfare. This would give 2 regional and 3 long-distance trains in the do-nothing scenario, and 3 regional trains and 5 long-distance trains in the investment scenario, yielding an investment benefit of 101 SEK/h. In any of these cases, the operator(s) can easily act strategically by stating intended timetables that increase investment benefits.

Table 5. An example with two types of trains (regional and long-distance).

<table>
<thead>
<tr>
<th>Without investment</th>
<th>With investment</th>
<th>Total appraisal benefits (kSEK)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R=regional trains, L = long-distance trains)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual outcome: the regional operator maximizes CS first. The long-distance operator maximizes PS, given number of regional trains.</td>
<td>3R,3L</td>
<td>6R,5L</td>
</tr>
<tr>
<td>Operators state that they will run the timetable that is socially optimal in the investment scenario.</td>
<td>3R,3L</td>
<td>3R,5L</td>
</tr>
<tr>
<td>Actual timetable outcome in the investment scenario applied in both the do-nothing and investment scenarios.</td>
<td>6R,5L</td>
<td>6R,5L</td>
</tr>
</tbody>
</table>

Examples of principles for choice of timetables

<table>
<thead>
<tr>
<th>Welfare maximizing</th>
<th>Profit maximizing</th>
</tr>
</thead>
<tbody>
<tr>
<td>2R,3L</td>
<td>3R,5L</td>
</tr>
<tr>
<td>1R,3L</td>
<td>2R,6L</td>
</tr>
</tbody>
</table>
4. Which principle should be used?

The above examples illustrate that the benefit of an investment is not defined without a principle defining timetables with and without the investment. The example also illustrates the problem of assuming timetables based on the stakeholder’s stated intentions. Even if it was possible to use binding contracts over 10-20 years, forcing operators to run the planned timetable in the forecast year, the problem of assuming a timetable for the do-nothing scenario remains. The same problem arises when there is a future target timetable, which is the case in some countries. The future target timetable is designed to be consistent with the investment scenario, but much less effort is usually spent on the timetable in the do-nothing scenario. Hence, the benefit of the investment is still arbitrary to some extent.

Carrying out sensitivity analyses using a range of different timetables could in some cases be seen as a solution, at least to illustrate the problem. Indeed, exploring the range of CBA outcomes resulting from different timetable assumptions is usually an enlightening exercise. However, doing and presenting several CBAs is often infeasible in practice because of limited time and resources. Moreover, if CBA results are to be used to compare different investment options against each other, they need to be well-defined: a result where several investment options have overlapping CBA ranges does not help a decision-maker.

Hence, an explicit principle for choosing timetables in the do-nothing and the investment scenarios is necessary to ensure comparability of different suggested investments. The timetables must be derived taking into account demand and supply as well as the prevailing institutional settings and capacity constraints. Like all CBA calculations, this principle should be transparent and possible to verify by a third party.

4.1 Current practice

The process of timetable construction in appraisal varies between countries using CBA for transport investment decisions, but there are also many similarities. To give an example, the current practice in Sweden can roughly be summarized as follows. Starting from current timetables, the plans and wishes from operators regarding traffic changes are collected by the national Transport Administration (the government agency responsible for railway investments and maintenance). Based on these stated plans and wishes, the Transport Administration identifies needs for capacity improvements. Based on a judgment of what capacity improvements are reasonable, operators’ stated plans are modified, and a future overall timetable is constructed, assuming these capacity improvements. This is (somewhat confusingly) called the “baseline timetable”. In most investment appraisals, the baseline timetable represents the investment scenario, while the do-nothing scenario is constructed by removing the investment, keeping the number of trains in the baseline timetable fixed, and increasing their travel times to meet capacity constraints.

Sometimes the current timetable is used in the do-nothing scenario and the planned timetable in the investment scenario. Although this seems like a better practice than to use the same timetable in the two scenario, this also opens up for the possibility for conscious or unconscious strategic behaviour from stakeholders wanting to promote certain investments and discredit others, perhaps in competition for national investment funding. The stakeholder may state (and perhaps honestly believe) that given the investment, the plan is to run a timetable that will maximize social benefits, although this timetable will not actually be the outcome because of
the prevailing institutional setting. This will also exaggerate benefits in the investment scenario since the current timetable is usually not socially optimal – especially as evaluated by the formal CBA framework. What appears in the appraisal to be “investment benefits” is actually the sum of the “proper” investments and merely moving to a more socially efficient timetable.

In current practice, it is seldom or never analysed how total benefits differ if the increased capacity is used to improve travel times, frequencies or reliability. Usually, it is up to the analyst to decide how the timetables should differ between the do-nothing and the investment scenarios. An example: during the preparation of the National Transport Investment Plan 2010-2021, 12 capacity-improving investments were analysed in the Stockholm area. The added capacity was used to improve frequencies in two cases, to improve travel times in three cases, to improve both frequency and travel times in three cases, and to improve reliability in four cases. (As to reliability improvements, there are no quantitative relationships at all, so the size of these effects was left to the analyst’s judgment.) In none of these cases was it analysed how benefits would have differed if the capacity improvement had been used differently.

4.2 Which principle?

It is clear that for CBA to serve its role as an impartial decision support, the principle for choosing appraisal timetables must be explicit, transparent and verifiable. What principle is most appropriate depends on the institutional context – the incentives and goals of various operators, and how scarce track capacity is allocated between operators. This varies between countries. Trains can be operated and/or owned by private companies, by national or by regional authorities. These may have different objectives, e.g. being profit-maximizing, welfare-maximizing or consumer surplus-maximizing, possibly within given budget constraints.

If there are no capacity constraints on the track, operators are free to choose the number of trains they wish to run. With knowledge about the goal of different operators, the analyst can in principle calculate (within the simplified framework of the CBA) the number of trains that each operator will run in the do-nothing and in the investment scenario. To accomplish this, a simplified model, including demand and supply such as the one set up in the present paper, taking into account the objectives of each operator must be set up.

If capacity is limited however, then the timetable principle must take into account how capacity is allocated between operators, given the prevailing institutional setting. Below we give some examples of different institutional contexts and discuss the setting up of a model for calculating the outcome timetables in the do-nothing and in the investment scenarios.

In countries where the railway is vertically integrated, infrastructure investments and timetable design are integrated decisions and there is only one operator. Hence, the allocation process and the resulting number of trains in the do-nothing and in the investments scenarios should be relatively easy to model given the objective of the operator. In settings like this the timetable in the investment scenario is probably well planned and close to the social optimum. Hence assuming social welfare-maximization in the do-nothing and in the investments scenario could also be an appropriate principle for designing appraisal timetables.

Countries where capacity is allocated through auctions or similar forms of market-based allocations (in the UK for instance), the model of the track allocation process for calculating actual outcome timetables in the do-nothing and in the
investment scenario must take the demand for tracks and the resulting prices into account.

In countries where the allocation process is a legal-administrative procedure, where a public agency determines how to the wishes of different operators should be negotiated (in Sweden, for example, where the Transport Administration determines which operators should get priority), it could be reasonable to assume that capacity is allocated to operators to maximize social welfare. Different operators have different objectives; long-distance trains and freight operators maximize profits, while regional train operators can be viewed as maximizing consumer surplus subject to a budget constraint. The Swedish track allocation process is far from transparent, however, even if it is supposed to strive for maximal social welfare. This principle has been difficult to operationalize, however, and the apparently operators differ in terms of bargaining power.

The principle of welfare maximization may also be the most appropriate approximation in many countries where the railway system is regulated and subsidised by society in several ways, including public provision or subsidisation of public transport, subsidised track access charges, public control over capacity allocation etc. The underlying idea would then be that the investment appraisal should capture the maximal social benefit that the infrastructure can generate and that society is trying to achieve maximal social welfare through its combination of regulation, subsidies and traffic provision. If, however, traffic supply is decided mainly by profit-maximizing train operators with substantial monopoly power, it would be natural to let the timetable be decided by maximal producer surplus. We argue that the essential thing is that there is an explicit model and principle for determining appraisal timetables.

4.3 Discrepancy with the realized timetable

There are always discrepancies between models and reality – in this case between reality and any CBA framework, with its cost and demand functions and its benefit valuations. A potential criticism of applying model-based principles for timetable construction in appraisal is that since the CBA framework is a simplified representation of reality, omitting or misrepresenting certain effects to some extent, the model/principle-based timetables will not necessarily be identical to the real current and later timetables. This is, however, not unique to the timetable assumptions, but to the entire cost-benefit analysis. When carrying out appraisals, we are working under the assumption that the CBA and model assumptions are sufficiently accurate representations of reality, in the sense that the CBA-ranking of investments or policy measure corresponds to the true ranking. Hence, even if the timetable computed according to the most appropriate principle, given the institutional setting, does not correspond to the actual outcome, it is necessary to stick to a single predefined principle.

If not sticking to a principle, this opens up a particularly insidious comparison: comparing the actual, current timetable (which can be observed) with the timetable that gives the highest benefits after the investment (in case the principle of social welfare maximization is applied). Few people would then argue against such a comparison – after all, it seems natural to use the actual timetable as a reference. But

---

6 Choosing profit maximization as the principle with which to construct timetables makes the practice to assume constant and exogenous fares even more suspicious than usually, so this should then be modified.
even if the current timetable is well-designed in reality (for the sake of argument), it may not be optimal when evaluated in the formal CBA framework. Then, if the analyst chooses a timetable in the investment scenario which is optimal when evaluated by the CBA framework – and why shouldn’t she? – benefits will be exaggerated. Even an investment that does not improve anything will show a benefit according to such a CBA. This shows that it is necessary to stick to the same CBA framework and timetable principle both in the investment and the do-nothing scenarios.

5. Conclusions

The social benefit of a railway investment cannot be defined without specifying timetables with and without the investment. Without an explicit principle for designing these timetables, there is a risk for arbitrariness, conscious or unconscious strategic behaviour when conducting the appraisal, counterintuitive appraisal outcomes, and a lack of comparability between suggested investments. This applies to all types of railway investments, but it is especially conspicuous and potentially insidious for capacity improvements, since the added capacity can be used in several ways – improved travel times, frequencies or reliability.

The conclusion that the timetables are important is in one sense almost trivial, but the awareness about the extent of the issue seems to be limited. Going through research literature and guidelines for railway appraisal, we have found no more than mentioning en passant of this issue or what principles should govern the design of timetables. Sometimes it is mentioned that it is important to have a “good” timetable for the planned investment to make the analysis reveal the highest possible benefit, but not that the timetable in the do-nothing scenario is equally important for the analysis, nor any discussion about whether this “good” timetable is likely to be realized.

We have illustrated how the appropriate timetable principle depends on the prevailing institutional setting – operator incentives and capacity allocation process. We argue that the most appropriate and general principle is to assume a simplified model of the track allocation process, taking into account capacity constraints, demand and supply, and assume the calculated actual outcome timetables in the do-nothing and in the investment scenario.

In some contexts, assuming socially optimal timetables might be a reasonable approximation of what society tries to achieve through regulations, policies, public transport provision, capacity allocation, pricing, subsidies etc. In other contexts, profit-maximizing timetables may be a more realistic approximation.

Whatever principle is chosen, a process is needed to systematically explore the social benefits of different timetables. To do this, current appraisal and modelling methods have to incorporate railway capacity analysis from the railway operations research tradition. At a minimum, one should explore how different timetables (in the do-nothing and the investment scenario) affect the total benefits of an investment.

Since the cost-benefit analysis is a simplified representation of reality, omitting or misrepresenting certain effects to some extent, the CBA-optimal timetable will not necessarily be the reality-optimal timetable. But when carrying out appraisals, we are working under the assumption that the CBA is a sufficiently accurate representation of reality, in the sense that the CBA-ranking of investments or policy measure corresponds to the true ranking. A similar assumption must be made for the evaluation of timetables. If it turns out that the assumed model of the track allocation
process is clearly unrealistic, this would mean that we are working with an unrealistic CBA framework, that is not a sufficiently accurate representation of reality to base judgments or ranking on.

Railway investments often represent huge public spending commitments, and hence it should be uncontroversial that decision-makers need reliable appraisal methods for comparing suggested investments against each other and against other uses of public resources. In this paper, we have argued that having an explicit, verifiable principle for timetable construction is necessary for CBA to be meaningful. Without such a principle, and a way of verifying that it is applied, CBA for railway investments will be a pointless exercise. In worst case, it becomes an instrument for lobbyist wolves dressed in the sheepish clothes of transport economists.

6. References


