Driver schedule efficiency vs. public transport robustness
A framework to quantify this trade-off based on passive data

Yap, Menno; van Oort, Niels

Publication date
2018

Document Version
Accepted author manuscript

Published in
Proceedings of CASPT 2018

Citation (APA)
Yap, M., & van Oort, N. (2018). Driver schedule efficiency vs. public transport robustness: A framework to quantify this trade-off based on passive data. In *Proceedings of CASPT 2018: 23-25 July, Brisbane, Australia* Article 5

CASPT 2018 Full Paper

Driver schedule efficiency vs. public transport robustness: A framework to quantify this trade-off based on passive data

Menno Yap · Niels van Oort

Abstract More complex, efficient driver schedules reduce operator costs during undisrupted operations, but increase the disruption impact for passengers and operator once a disruption occurs. We develop an integrated framework to quantify the passenger and operator costs of disruptions explicitly as function of different driver schedule schemes. Since the trade-off between driver schedule efficiency and robustness can be quantified, this supports operators in their decision-making.

Keywords: Disruptions · Driver scheduling · Passenger perspective · Passive data · Public Transport · Robustness

1 Introduction

The Driver Scheduling Problem (DSP) for public transport networks is a well-studied topic in operations research (e.g. Kroon and Fischetti, 2001; Huisman et al. 2005; Portugal et al. 2009; De Leone and Festa, 2011). Research developments and the availability of advanced driver scheduling software (such as HASTUS) have resulted in the development and implementation of more complex driver schedules, which can improve operator efficiency and reduce operating costs. Where a drivers' duty traditionally consisted of tasks on one vehicle only, a duty now often consists of tasks on different vehicles during one shift. This is implemented by 'single-line multi-vehicle' scheduling – a driver changes vehicles during a shift, but remains operating one and the same line – or 'multi-line multi-vehicle' scheduling. In the latter case, a more complicated driver schedule is applied where driver tasks are scheduled on different vehicles as well as on different lines during one shift. This allows vehicles to operate with a different driver during a drivers' break, and can reduce the total required fleet size and number of driver hours the operator requires on a network level.

Menno Yap
Department of Transport and Planning, Delft University of Technology
Delft, the Netherlands
Email: M.D.Yap@TUDelft.nl

Niels van Oort
Department of Transport and Planning, Delft University of Technology / Goudappel Coffeng consultants
Delft, the Netherlands
Email: N.vanOort@TUDelft.nl

There is however a trade-off between driver schedule complexity and public transport robustness: albeit a more complex driver schedule reduces operator costs during undisrupted operations, it increases the impact of disruptions for passengers and operator in case a disruption occurs. More complex schedules result in longer service recovery times once the incident has been resolved, and in more complicated and less effective rescheduling: there is a risk of delay propagation over the network if a driver is not able to arrive in time for the next task on another line.

Robust driver scheduling studies mainly incorporate minor recurrent delays by adding slack in the timetable (e.g. Laplagne 2008), but do not consider robustness related to large non-recurrent disruptions. On the other hand, studies aiming to quantify the passenger impact of urban public transport disruptions (e.g. Van Oort et al. 2015b; Jenelius and Cats 2015; Cats et al. 2016, Yap et al. 2018c) do not incorporate driver schedule complexity. In our study we develop an integrated framework in which passenger disruption impact is explicitly quantified as function of different levels of driver schedule complexity. For operators to balance schedule efficiency and robustness, quantification of the operator impact of disruptions for different types of driver schedules is explicitly incorporated in this framework.

2 Methodology

Table 1 shows the notations used in our framework.

Table 1 Indices and sets, parameters and variables

Indices and sets	
s, S	stop index, set
l, L	line index, set
S_l	set of stops on line l, $S_l \subseteq S$
$l = \{s_{l,1}, s_{l,2}, \dots, s_{l,\lvert l \rvert}\}$	line l is defined as ordered sequence of stops
r, R	run index, set
R_l	set of runs on line l, $R_l \subseteq R$
i	index for disruption
h	hourly time period
Parameters	
β_1	weight of perceived passenger waiting time
β_2	operator revenue for average passenger journey
β_3	operator costs for each hour of personnel overtime
β_4	operator fine for run with too early departure
β_5	operator fine for run with too late departure
β_6	operator fine for cancelled run
β_7	operator fine for unavailable infrastructure per hour
E_d	demand elasticity
VoT	Value-of-Time
γ	crowding in-vehicle time multiplier
γ^s	crowding in-vehicle time multiplier at seat capacity
γ^c	crowding in-vehicle time multiplier at crush capacity
φ_r^s	seat capacity of run r
φ_r^c	crush capacity of run r

Variables	
$\tilde{t}_{rs}^{a}$	scheduled arrival time of run r at stop s
$\tilde{t}_{rs}^{d}$	scheduled departure time of run r from stop s
t_{rs}^{a}	arrival time of run r at stop s
t_{rs}^{d}	departure time of run r from stop s
$t_{rs_l}^{ivt}$	passenger in-vehicle time of run r from stop s_l to s_{l+1}
$t_{rs_l}^{ivt,p}$	perceived passenger in-vehicle time of run r from stop s_l to s_{l+1}
t_{s}^{wtt}	passenger waiting time at stop s
$t_{s_i s_j}^{p}$	generalized passenger travel time for journey from stop i to stop j
t^{i}	duration of disruption
t^{o}	personnel overtime hours per disruption
c_{o}^{i}	operator costs of disruption
f_{l}^{h}	frequency of line l during hour h
d_{r}	headway between run r and subsequent run r^{+}
q_{rs}	passenger load on-board run r between stop s and subsequent stop
q_{rs}^{in}	number of passengers boarding run r at stop s

To quantify the costs of a public transport disruption we develop a framework as shown in Figure 1.

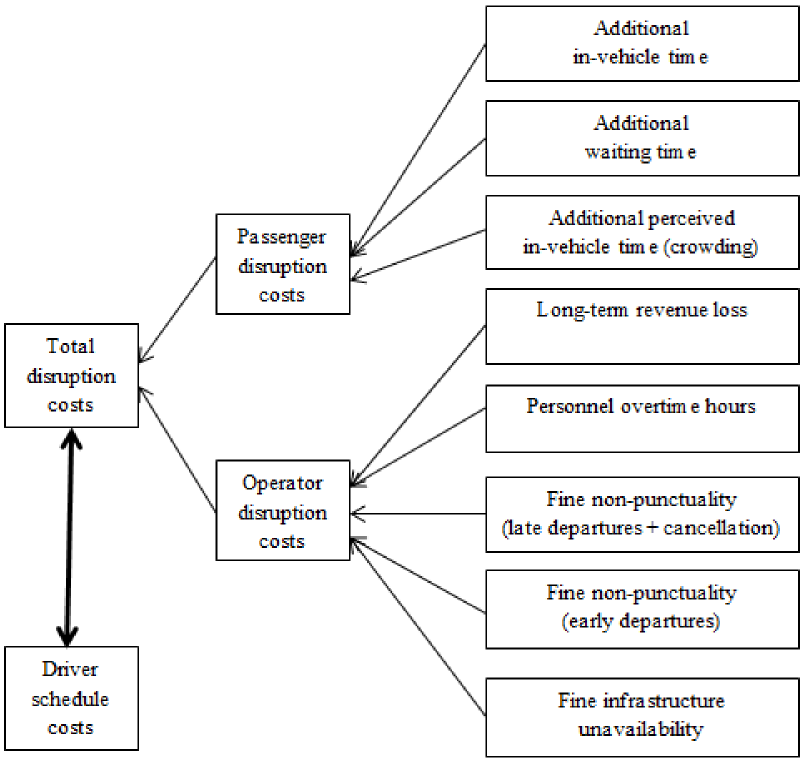

Fig. 1 Framework to quantify disruption costs compared to driver schedule costs

Costs are divided into passenger costs and operator costs, which both consist of several components. The (societal) disruption costs for passengers consist of the additional in-vehicle time, waiting time and perceived in-vehicle time due to crowding, all expressed in monetary terms. The additional in-vehicle time equals the delay of each run $r \in R$ due to this disruption, multiplied by the passenger flow q_{rs} travelling over the disrupted link between s_l and s_{l+1} (Eq.1).

$$\Delta\, t^{ivt} = \sum_{r \in R} \left(\left(\left(t^{a}_{rs_{l+1}} - \tilde{t}^{a}_{rs_{l+1}} \right) - \left(t^{d}_{rs_l} - \tilde{t}^{d}_{rs_l} \right) \right) * q_{rs_l} \right) * VoT \tag{1}$$

The additional waiting time is quantified by comparing the scheduled and realized headway, incorporating the Percentage Regularity Deviation Mean (PRDM) as measure for irregularity for each service hour (Van Oort & Van Nes 2009) (Eq.2). Given our focus on urban, high frequent public transport services, a random passenger arrival pattern is assumed resulting in the quantification of additional waiting time due to irregularity as shown in Eq.3.

$$PRDM^h = \frac{\sum_{r^h \in R^h} \left| \frac{d_r^h - \tilde{d}_r^h}{\tilde{d}_r^h} \right|}{\frac{60}{2 * f^h}} \tag{2}$$

$$\Delta\, t^{wtt} = \sum_{h \in H} \left(\left(\frac{60}{2 * f_l^h} \right) * \left(1 + (PRDM^{h^2}) \right) - \left(\frac{60}{2 * \tilde{f}_l^h} \right) \right) * \beta_1 * VoT \tag{3}$$

Since large disruptions can result in service cancellations and more irregular service headways, the average crowding level on the remaining runs is expected to increase. As crowding results in a higher perceived in-vehicle time, this component is quantified as well (Eq.4). For the public transport lines directly affected by the disruption, as well as parallel lines used as alternative route by passengers, for each run and each link the average crowding level is compared between an average undisrupted day and during the disruption, thereby correcting for seasonal effects and day of the week. Based on the vehicle seat capacity φ_r^s and crush capacity φ_r^c and their corresponding crowding multipliers γ_r^s and γ_r^c, the realized in-vehicle time is multiplied by a crowding multiplier γ_{rs}. In line with e.g. Wardman and Whelan (2011) and Yap et al. (2018a), γ_{rs} is assumed to be a linear piecewise function between 50% seat occupancy, seat capacity and crush capacity (Eq.5).

$$\Delta\, t^{ivt,p} = \sum_{r^h \in R^h} \sum_{s_{l,1} \in s_{l,|l|}} \left((q^{i}_{rs} * (t^{a}_{rs+1} - t^{d}_{rs}) * \gamma_{rs} \right) - \left(q^{j \neq i}_{rs} * (t^{a}_{rs+1} - t^{d}_{rs}) * \gamma_{rs} \right) * VoT \tag{4}$$

$$\gamma_{rs} = \begin{cases} 0.95 & if\ q_{rs} \leq 0.5 * \varphi_r^s \\ 0.95 + \left(\frac{q_{rs} - 0.5 * \varphi_r^s}{0.5 * \varphi_r^s}\right) * (\gamma_r^s - 0.95) & if 0.5 * \varphi_r^s < q_{rs} < \varphi_r^s \\ \gamma_r^s + \left(\frac{q_{rs} - \varphi_r^s}{\varphi_r^c - \varphi_r^s}\right) * (\gamma_r^c - \gamma_r^s) & if\ q_{rs} > \varphi_r^s \end{cases} \tag{5}$$

One component of operator disruption costs is the lost revenues following a loss of public transport demand due to the impact of disruptions. Although long-term ridership impacts from disruptions are difficult to predict, we used a simple elasticity-based approach as applied by Van Oort et al. (2015a), using parameters calibrated for planned disruptions based on smart card data (Yap et al. 2018b). For a given time period T, the generalized travel time is calculated for the disrupted and undisrupted scenario (Eq.6). The generalized costs equal the weighted sum for the disrupted scenario i and undisrupted scenario $j \neq i$, as ratio of the duration of a disruption t^i compared to T, and is compared to a scenario with no disruptions during T (Eq.7).

$$\bar{t}^p = \frac{\sum_{s_i \in S_i} \sum_{s_j \in S_j} \left(\left(t_{s_i}^{wtt} * \beta_1 + t_{s_i,s_j}^{ivt,p} \right) * q_{s_i,s_j} \right)}{\sum_{s_i \in S_i} \sum_{s_j \in S_j} q_{s_i,s_j}} \tag{6}$$

$$\Delta q = \left(E_d * \left(\frac{\bar{t}^{p i} * t^i + \left(\bar{t}^{p j \neq i} * (T - t^i) \right)}{\bar{t}^{p j \neq i} * T} - 1 \right) + 1 \right) * \sum_{s_i \in S_i} \sum_{s_j \in S_j} q_{s_i,s_j} \tag{7}$$

The demand loss is quantified in Eq.8 by multiplication of Δq with the average passenger revenue. Due to the unannounced and relatively heavy impact of unplanned disruptions compared to planned disruptions, this cost component can be considered a lower bound. For each disruption the extra overtime hours for personnel, the number of early runs (departure before scheduled departure time), late runs (departure later than scheduled departure time plus threshold Δ) or cancelled runs, and the time the infrastructure is not available, are multiplied with their corresponding cost parameters (Eq.8). For the latter four components, the values of the cost parameters are usually specified in the contract between operator and authority, indicating the fine for each early, late or cancelled run, or for each hour that no PT services can be provided on a link resulting from infrastructure unavailability.

$$c_o^i = \beta_2 * \Delta q + \beta_3 * t + \beta_4 * \sum_{r \in R} r^e + \beta_5 * \sum_{r \in R} r^l + \beta_6 * \sum_{r \in R} r^c + \beta_7 * t^i$$

$$with\ r^e \begin{cases} 1\ if t_{rs}^d < \tilde{t}_{rs}^d \\ 0\ if\ t_{rs}^d \geq \tilde{t}_{rs}^d \end{cases},\ r^e \begin{cases} 1\ if t_{rs}^d > \tilde{t}_{rs}^d + \Delta \\ 0\ if\ t_{rs}^d \leq \tilde{t}_{rs}^d + \Delta \end{cases},\ r^c \begin{cases} 1\ if run\ has\ been\ cancelled \\ 0\ if\ run\ is\ not\ cancelled \end{cases} \tag{8}$$

3 Case study

We apply our framework to the urban public transport network of The Hague, the Netherlands, which consists of 12 light rail / tram lines and 8 urban bus lines. One large disruptions on the light rail track is considered as case study (Figure 2), which occurred Wednesday January 6th, 2016 due to a switch failure between 11:22h and 14:33h. At 19:38h all services were running according to schedule again. The disruption resulted in splitting light rail services 3 and 4, normally operating between the city of The Hague and the satellite city Zoetermeer (Figure 2 purple and orange, respectively) in a western and eastern part and cancellation of some services due to turning capacity constraints.

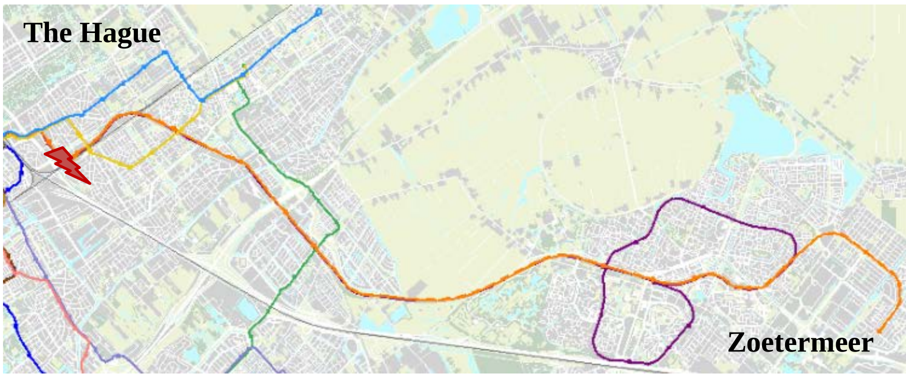

Fig. 2 Urban public transport case study network The Hague

To demonstrate our proposed framework, we compare the disruption costs and driver schedule costs for two different driver schedule scenarios for this disruption.

Scenario 1: multi-line multi-vehicle schedule with punctuality-based rescheduling
This scenario describes the situation as currently applied by the operator of the case study, namely applying a multi-line multi-vehicle driver schedule. A punctuality-based rescheduling approach is applied, aiming to let the remaining services depart according to schedule where possible. Although headway-based control is preferred from a passenger perspective, interviews with public transport controllers indicate that the complexity of the multi-line multi-vehicle schedule requires punctuality-based control. By trying to keep departure times of remaining services close to schedule, delay propagation to other lines – resulting from drivers arriving earlier or later than scheduled for the next task of their shift on another line – is aimed to be reduced. The operator does not use rescheduling software which allows for headway-based control when this relatively complex multi-line multi-vehicle schedule is applied.

The components of our framework related to passenger disruption costs are quantified directly using realized Automated Fare Collection (AFC) and Automated Vehicle Location (AVL) data for the Wednesday the disruption occurred, and three other Wednesdays of the same month without disruptions, so that particularly the disruption impact on crowding can be compared to regular, undisrupted days. AVL data is also used to quantify the number of early, late and cancelled services. Based on AVL data, log-files and information provided by public transport schedulers, the number of personnel overtime hours and the time the infrastructure was unavailable for PT services are determined for this scenario. Parameter values for operator fines β_3 to β_7 are determined from the contractual agreements between PT operator and authority.

Scenario 2 single-line multi-vehicle schedule with headway-based rescheduling
We contrast the disruption costs of scenario 1 with scenario 2, a scenario which evaluates the passenger and operator disruption costs in case a single-line multi-vehicle driver schedule would be applied. In this case, drivers only shift between vehicles of the same line during one duty. This has two effects. First, headway-based control can be applied to remaining services, since there is no risk of delay propagation to other lines (HTM, 2015). Second, this less complex driver schedule reduces the recovery time of PT services from the disruption, which reduces both the passenger disruption costs, and the personnel overtime hours.

Since this scenario is currently not applied by the case study operator, the disruption costs cannot be inferred directly from AFC, AVL and log-data in this case. Values for this scenario can be obtained by combining quantitative and qualitative sources. Based on realized AFC and AVL data when applying punctuality-based control, we can simulate the disruption impact on passenger in-vehicle time, waiting time and crowding if all remaining services would be supplied with an equal headway in case of headway-based control. When applying our framework, headway-based control affects the additional waiting time. We calculated the PRDM for an average undisrupted day based on AVL data (which equals 0.2 for our case study services), and constrained the PRDM for each disruption hour to this value to quantify the reduced additional waiting time. Based on the remaining services and the PRDM being capped at a value of 0.2, the perceived service frequency can be calculated (Van Oort and Van Nes 2009). By dividing the hourly passenger load equally by the perceived service frequency, the expected occupancy for each run is calculated, resulting in monetized additional perceived in-vehicle time due to crowding for this scenario. The generalized travel time during disruptions is updated as consequence, adjusting the expected revenue loss from demand reduction. Personnel overtime hours are expected to decrease linearly with the service recovery time reduction (HTM, 2015). Based on calculations of the impact of different driver schedule types on service recovery time performed by the case study operator, and interviews held with public transport schedulers and controllers, the service recovery time is expected to reduce by ≈50% (De Bont and Wagemans, 2015). This allows quantification of the reduced costs from personnel overtime, as well as the shortened passenger impact of the disruption. Services are now assumed to operate according to an undisrupted day 2.5 hours after the disruption has been resolved (at 17:00h), instead of the service recovery time of 5 hours which is currently the case.

Multiplication of the disruption costs by the yearly number of disruptions based on log-data allows for the quantification of yearly passenger and operator costs for different driver schedule scenarios. The reduced disruption costs resulting from a less complex driver schedule can then be compared to the increased driver schedule costs, so that the trade-off between disruption and schedule costs can be monetized.

4 Results

4.1 Results

From Figure 3 we can conclude that one non-recurrent disruption on the considered light rail network currently (scenario 1: multi-line multi-vehicle scheduling with punctuality-based control) costs ≈€65,000, consisting of ≈€36,000 passenger costs and ≈€29,000 operator costs. The additional waiting time costs and long-term revenue loss are the most important cost components.

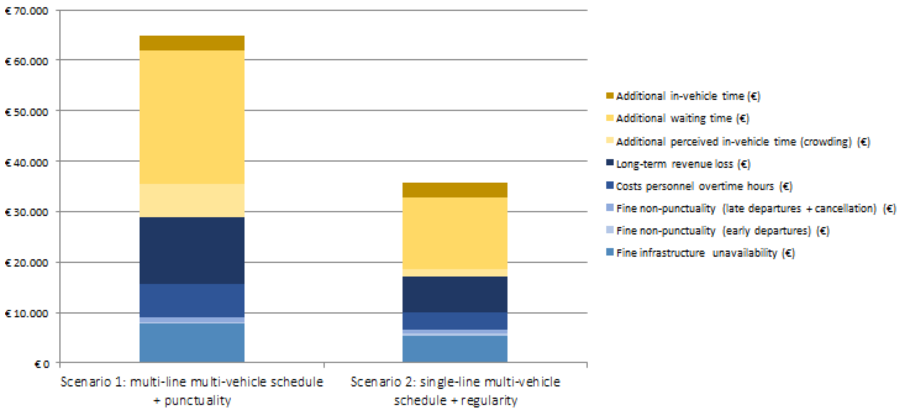

Fig. 3 Costs per disruption per component for different driver schedule types

When scenario 2 – single-line multi-vehicle scheduling with headway-based control – would be applied, total disruption costs are expected to decrease by 45% to ≈€36,000 per disruption. This is especially caused by less additional waiting time and lower additional perceived in-vehicle time, due to the improved regularity between services and shorter service recovery times. This, in turn, reduces revenue losses from long-term passenger demand decrease. When extrapolating these costs to yearly costs based on the frequency of non-recurrent disruptions, one can conclude from Figure 4 that yearly disruption costs are expected to be equal to ≈€1.1 million and ≈€0.6 million for scenario 1 and scenario 2, respectively.

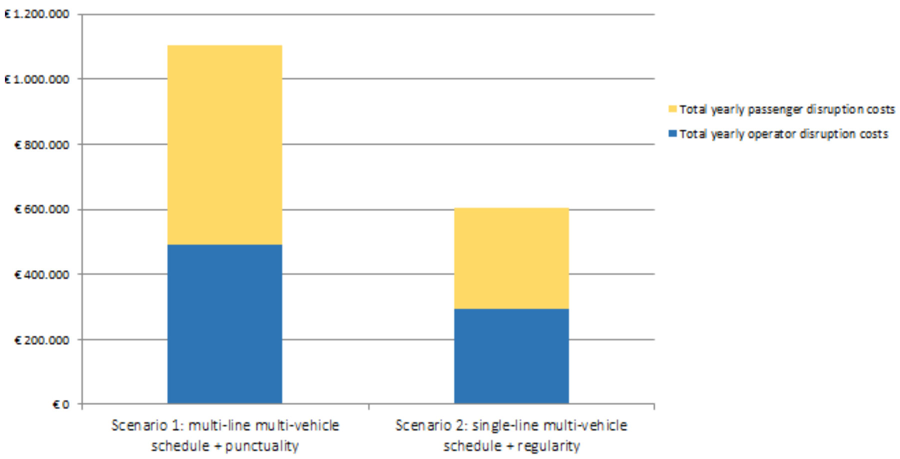

Fig. 4 Yearly passenger and operator disruption costs for different driver schedule types

In Figure 5 the trade-off between disruption costs and driver schedule costs is quantified for single-line multi-vehicle scheduling (scenario 2) compared to the current multi-line multi-vehicle scheduling (scenario 1) applied to the case study network. A less complex and less efficient driver schedule without shifts between different lines increases the direct driver schedule costs by €300,000 (HTM, 2015), but reduces the total disruption costs by €500,000 and is beneficial from a societal perspective. The operator disruption costs are reduced by €200,000, showing that purely the financial robustness benefits of this less complex driver schedule do not outweigh the costs.

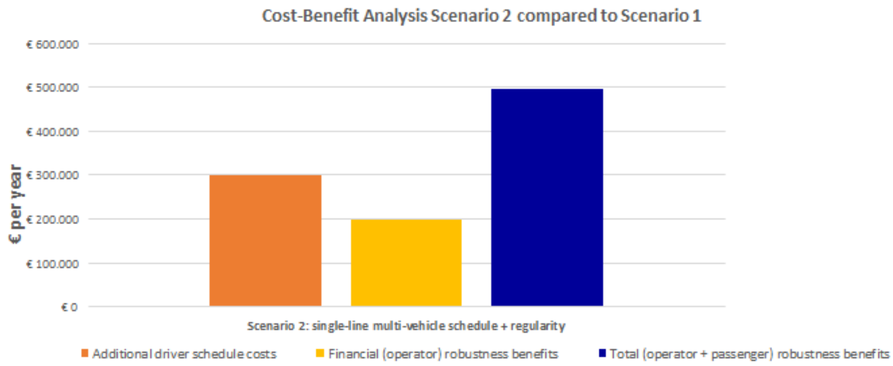

Fig. 5 Cost-Benefit Analysis for trade-off between disruption and driver schedule costs

4.2 Sensitivity analysis

A sensitivity analysis is performed to the two most uncertain parameters: the demand elasticity and the impact of single-line multi-vehicle scheduling on service recovery time reduction. We experimented with values of -0.3 and -0.7 for demand elasticity, compared to the default value of -0.5 [-40%,+40%]. For the reduction in service recovery time, a reduction of 30% and 70% was tested next to the default value of 50% [-40%,+40%].

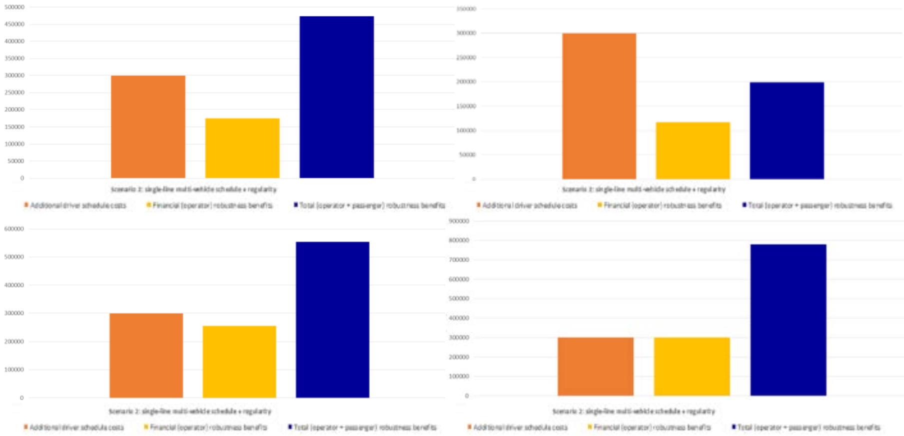

Fig. 6 Sensitivity analysis to demand elasticity (-0.3: upper left / -0.7 (lower left) and service recovery time reduction (30%: upper right / 70%: lower right)

Figure 6 shows that a 40% less negative demand elasticity parameter of -0.3 reduces the operator robustness benefits of Scenario 2 by €50,000, showing a relatively limited sensitivity of the outputs to this parameter value. If service recovery time reduction is 40% less than assumed, operator robustness benefits reduce by almost €100,000, whereas total robustness benefits reduce by €300,000. Results show to be especially sensitive to this parameter, indicating that more in-depth research to this value is recommended.

5 Conclusions

In this study we develop a framework to quantify the passenger and operator costs of disruptions explicitly as function of different driver schedule schemes. This supports operators in their decision-making, since the trade-off between driver schedule complexity and efficiency on the one hand, and robustness on the other hand, can be quantified. We test our proposed framework for one large, non-recurrent disruption on the case study network of The Hague, the Netherlands. Results for this case study show that when applying a less complex, single-line multi-vehicle driver schedule, total monetized passenger and operator robustness benefits outweigh the increased driver schedule costs. The financial robustness benefits for the operator solely are however smaller than the increased operator costs resulting from a less efficient driver schedule. We recommend particularly more in-depth research to the impact of different types of driver schedules on (the reduction of) service recovery time from a disruption.

References

Cats, O., Yap, M.D., Van Oort, N. (2016). Exposing the role of exposure: Public transport network risk analysis. Transportation Research Part A, 88, 1-14.

De Bont, R., Wagemans, T. (2015). Slippen misschien toch niet onmogelijk voor (Regie) Bijsturing. Technical report (in Dutch). The Hague, the Netherlands: HTM Personenvervoer NV.

De Leone, R., Festa, P. (2011). A Bus Driver Scheduling Problem: a new mathematical model and a GRASP approximate solution. Journal of Heuristics, 17(4), 441-466.

HTM Personenvervoer NV (2015). Expert meeting with controller T. Wagemans, personnel manager S. Rijke and planning manager G. Keuzenkamp, 16th December 2015.

Huisman, D., Freling, R., Wagelmans, A.P.M. (2005). Multiple-depot integrated vehicle and crew scheduling. Transportation Science, 39(4), 491–502.

Jenelius, E., Cats, O. (2015). The value of new public transport links for network robustness and redundancy. Transportmetrica A, 11(9), 819-835.

Kroon, L., Fischetti, M. (2001). Crew scheduling for Netherlands railways destination: customer. In: Voß, S., Daduna, J.R. (eds.) Computer-Aided Scheduling of Public Transport, pp. 181–201. Springer, Berlin (2001).

Laplagne, I.E. (2008). Train Driver Scheduling with Window of Relief Opportunities. University of Leeds: PhD Thesis.

Portugal, R., Lourenço, H., Paixão, J. (2009). Driver scheduling problem modelling. Public Transport, 1, 103–120.

Van Oort, N., Van Nes, R. (2009). Regularity analysis for optimizing urban transit network design. Public Transport, 1(2), 155-168.

Van Oort, N., Brands, T., De Romph, E. (2015a). Short-term prediction of ridership on public transport with smart card data. Transportation Research Record, 2535, 105-111.

Van Oort, N., Brands, T., De Romph, E., Aceves Flores, J. (2015b). Unreliability Effects in Public Transport Modelling. International Journal of Transportation, 3(1), 113-130.

Wardman, M., Whelan, G. (2011). Twenty years of rail crowding valuation studies: evidence and lessons from British experience. Transport Reviews, 31, 379-398.

Yap, M.D., Cats, O., Van Arem, B. (2018a). Crowding valuation in urban tram and bus transportation based on smart card data. Transportmetrica A, under review.

Yap, M.D., Nijenstein, S., Van Oort, N. (2018b). Improving predictions of public transport usage during disturbances based on smart card data. Transport Policy, 61, 84-95.

Yap, M.D., Van Oort, N., Van Nes, N., Van Arem, B. (2018c). Identification and quantification of link vulnerability in multi-level public transport networks: a passenger perspective. Transportation, DOI: 10.1007/s11116-018-9892-5.