Waiting times in Copenhagen Airport

Estimation and evaluation on new data

April 2017

Danmarks Tekniske Universitet

Waiting times in Copenhagen Airport

Report [Nr.]
2017

By

Cover photo: [Text]
Published by: Transport DTU, Building 116V, DK-2800 Kgs. Lyngby, Denmark
Request report from: www.transport.dtu.dk
ISSN: [0000-0000] (electronic version)
ISBN: [000-00-0000-000-0] (electronic version)

ISSN: [0000-0000] (printed version)
ISBN: [000-00-0000-000-0] (printed version)

Content

Summary

The current report concerns phase 3 of a project that looks at waiting times in the central security check in Copenhagen Airport. Phase 1 built a prediction model for the waiting times and used that to perform an economic evaluation of increasing the capacity in the security check. Phase 2 analysed the consequences of a data issue that was discovered. The purpose of the current phase 3 is to update phase 1 with the most recent data, in which the issue analysed in phase 2 has been resolved.

This report evaluates the cost to travellers of waiting time in the central security check in Copenhagen Airport. It predicts the reduction in passenger cost that follows from increasing the capacity in the central security check and compares this reduction to the cost to the airport of increasing the capacity. The general conclusion is that considerable reductions in waiting time costs are feasible and that the savings for passengers outweigh the additional cost to the airport over a large range of reductions. Requiring the airport to reduce waiting times in the central security check will therefore yield a net benefit to society.

The physical capacity in the central security check is 20 lanes. The number of lanes that are manned and open varies over the day. The present analysis considers only the opening of more of the 20 lanes, while the physical capacity is retained as it is now.

We have used data that describe the actual waiting times, the number of open lanes, and the number of passengers passing security. We have observations every 15 minutes over a period of about 4 months, which means that the present analysis is based on a substantial database: altogether we have 8,170 observations.

Based on these data, we have developed a statistical model that predicts the waiting time cost to travellers as it depends on the number of open lanes. We find that the statistical model gives a satisfactory description of the historical data. We have ensured that our estimates of the cost reductions that follow a capacity increase are conservative: they will tend to be on the low side of the actual cost reductions that may be achieved. The model is strongest in the range where we have most data. In the presentation of results we omit the times outside the interval from 6am to 8pm every day where the number of passengers and the number of open lanes are low and the model predictions therefore are less reliable. This increases our confidence in the predictions that we present.

The passenger waiting time cost depends on the mean waiting time and on the random variability of waiting time. This takes into account that it is not only the waiting time that matters to passengers but also the uncertainty they face regarding how long the waiting time will be when they arrive at the airport.

We have simulated four scenarios describing an average week in 15 minute intervals during the period from 6am to 8pm every day. A base scenario replicates the average week with the historical average number of open lanes every 15 minutes.

Three policy scenarios predict the consequences of opening 1, 2, and 3 additional lanes, respectively, at all times through each day, while staying within the physical capacity of 20 lanes. The distribution of waiting times is shown in Figure 1 for the base scenario and for the scenario with 3 additional lanes. Figure 2 shows for an average Wednesday how waiting times are affected across the day by adding 3 lanes.

We compare the cost savings to passengers from opening more lanes to the cost to the airport as estimated by the Danish Transport, Construction and Housing Agency. Both the modelling and the cost estimate are conditional on the existing physical lane capacity. We find that adding 1 or 2 lanes yields a net benefit at all times during the week. Adding three lanes yields a substantial net benefit in general, but there are a few 15 minute intervals during the week where the net benefit of the third lane becomes negative.

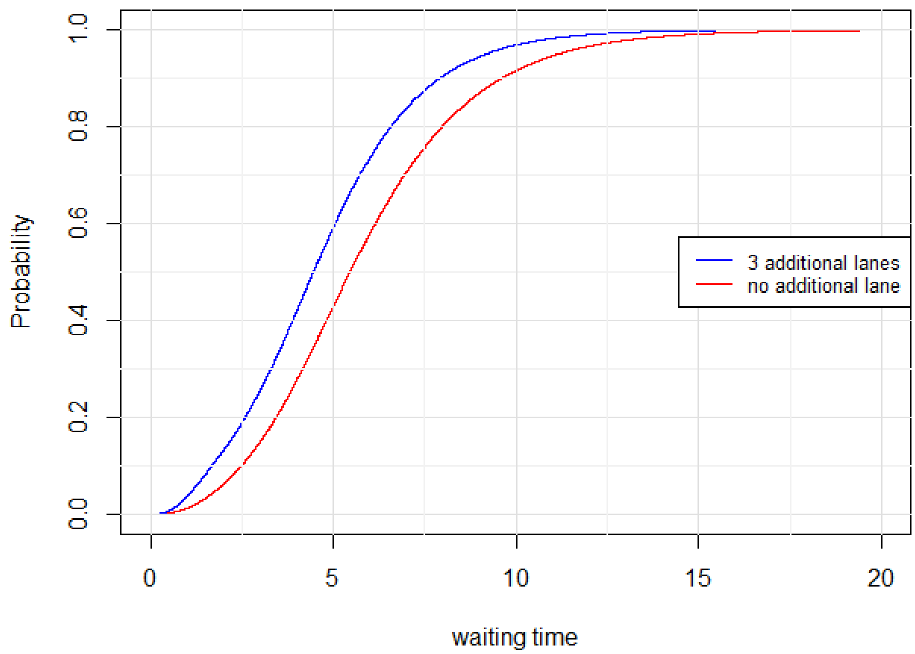

Figure 1 The distribution of waiting times in the base scenario and with 3 additional lanes

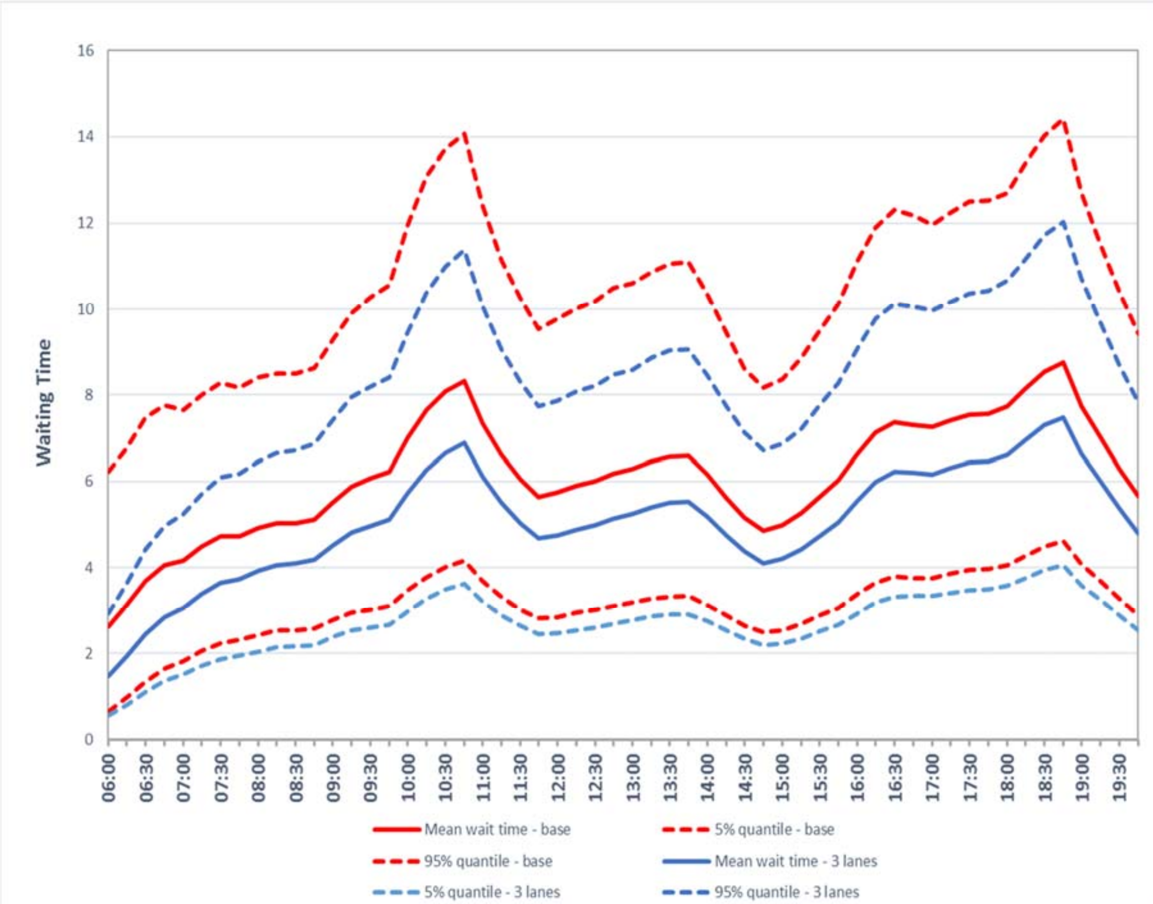

Figure 2 Waiting time statistics on an average Wednesday

Opening one additional lane yields an average benefit to passengers of 4947 DKK per hour with a corresponding cost to the airport of 1452 DKK per hour. Opening a second additional lane yields an additional average benefit to passengers of 3981 DKK per hour, which is still larger than 1452 DKK per hour. Opening a third additional lane yields an additional average benefit to passengers of 3338 DKK per hour, which is also larger than the cost to the airport of 1452 DKK per hour. These savings are larger than those computed in phase 1.

The uncertainty inherent in the model predictions increases as we add lanes and move away from the range we observe in the data. We therefore do not investigate further capacity increases as we would then be less confident about the model predictions. The implication is that the net benefits are so large that we are not confidently able to identify the break-even point where additional capacity no longer yields a net benefit.

In conclusion, the calculations indicate a clear net benefit of opening three additional lanes at all times during the day. The main points relevant for assessing the robustness of the conclusion are the following.

The benefits are proportional to the values of time and reliability. The values used are on the low side of the available evidence, which indicates that the actual benefit of a capacity increase is likely to be larger than the calculated benefit.

The cost per lane hour is directly proportional to the estimate from the Danish Transport, Construction and Housing Agency. The cost per lane hour would thus need to be more than 100% larger in order to change the conclusion.

We therefore find that the conclusion that there is a clear net benefit of opening three additional lanes at all times during the day is quite robust.

1. Background and purpose

The current report concerns phase 3 of a project that looks at waiting times in the central security check in Copenhagen Airport. Phase 1 built a prediction model for the waiting times and used that to perform an economic evaluation of increasing the capacity in the security check. Phase 2 analysed the consequences of a data issue that was discovered. The purpose of the current phase 3 is to update phase 1 with the most recent data, in which the issue analysed in phase 2 has been resolved.

Because of censoring issues in the measurement of waiting times at the central security check in Copenhagen Airport, it was decided to reestimate the model for waiting times found in phase 1 and use this new model to reevaluate the effect of capacity on waiting times.

This note describes this phase 3 of the project, which includes

- Management and checking of the new data from September 2016 to December 2016.
- A descriptive analysis of waiting times, demand and capacity in the central security check at Copenhagen Airport for data from September 2016 to December 2017.
- Reestimation of the waiting time model on these new data to assess the effect of capacity on the waiting time.
- Evaluation of the effect of extra capacity on the waiting time from a societal point of view.

The project is carried out by Mogens Fosgerau, Abhishek Ranjan and Stefan L. Mabit for the Ministry of Transport, Building and Housing.

2. Data and descriptives

2.1 Data

We have received the following data from the Danish Transport, Construction and Housing Agency.

- Waiting times, 15 minute bins, from September 1st, 2016, to December 31st, 2016.
- Number of lanes open, 15 minute bins, from September 1st, 2016, to December 31st, 2016.
- Number of passengers departing by each airplane from September 1st, 2016, to December 31st, 2016.

The latter file provides us with the number of passengers that departs from Copenhagen Airport. We use the time stamp in the file to aggregate them into 15 minute bins. The time stamp is in UTC. We have converted it to Danish time, taking daylight time saving into account.

We do not have the exact number of passengers passing the central security check (CSC) each 15 min but we do have information about the total number of travellers that actually used the CSC on a given day. We find this share to be 0.777 of the departing passengers as described in section 2.2. This is used to scale down the number of passengers on departing planes to the number of passengers that used CSC, i.e. excluding transit, fast track, etc.

2.2 Descriptives

The following figures uses the data from phase 3 where demand, number of open lanes, waiting times are available, i.e. the period September 1st, 2016, to December 31st, 2016. Figure 3 shows the average number of open lanes across the average day in 15 minute intervals (blue curve) and the average number of passengers embarking planes (red curve). The time stamps on the demand is the actual take-off time, which should occur some time after passengers pass through security. This issue of different time stamps is dealt with in the modelling, see section 3.

The number of passengers per 15 minutes fluctuates between approximately 300 and 700 persons from 6 am to 10 pm. Not all of these pass through the central security since there are fast track and other additional lanes. Based on the total daily number of passengers passing through security and the total daily number of passengers departing from the airport, we have calculated that an average share of 0.777 of the departing passengers pass through the CSC.

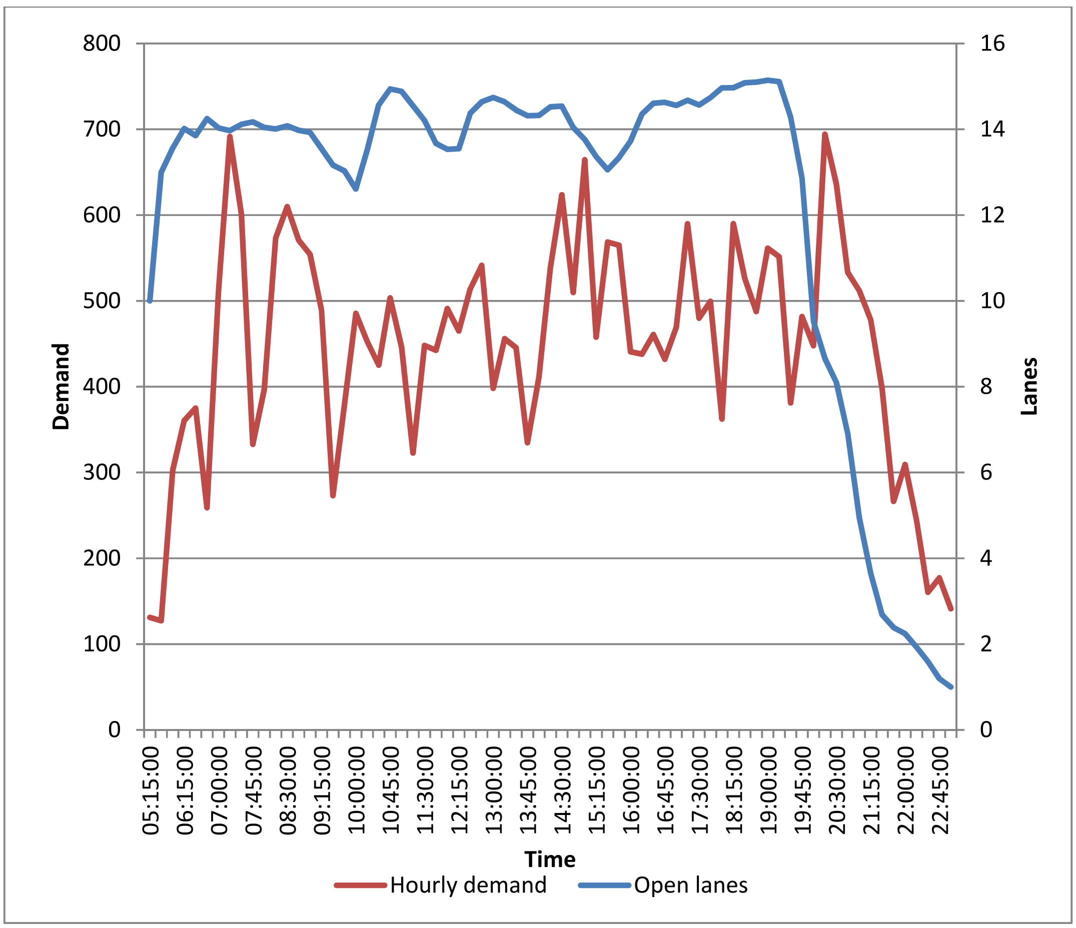

Figure 3 Average hourly demand and open lanes

Figure 4 shows the average waiting time in CSC (red curve) and the average number of open lanes (blue curve) for every 15 minute interval. The waiting time graph shows an expected morning peak as well as an late afternoon peak. But is also shows some variation in the middle of the day not related to the peak hours.

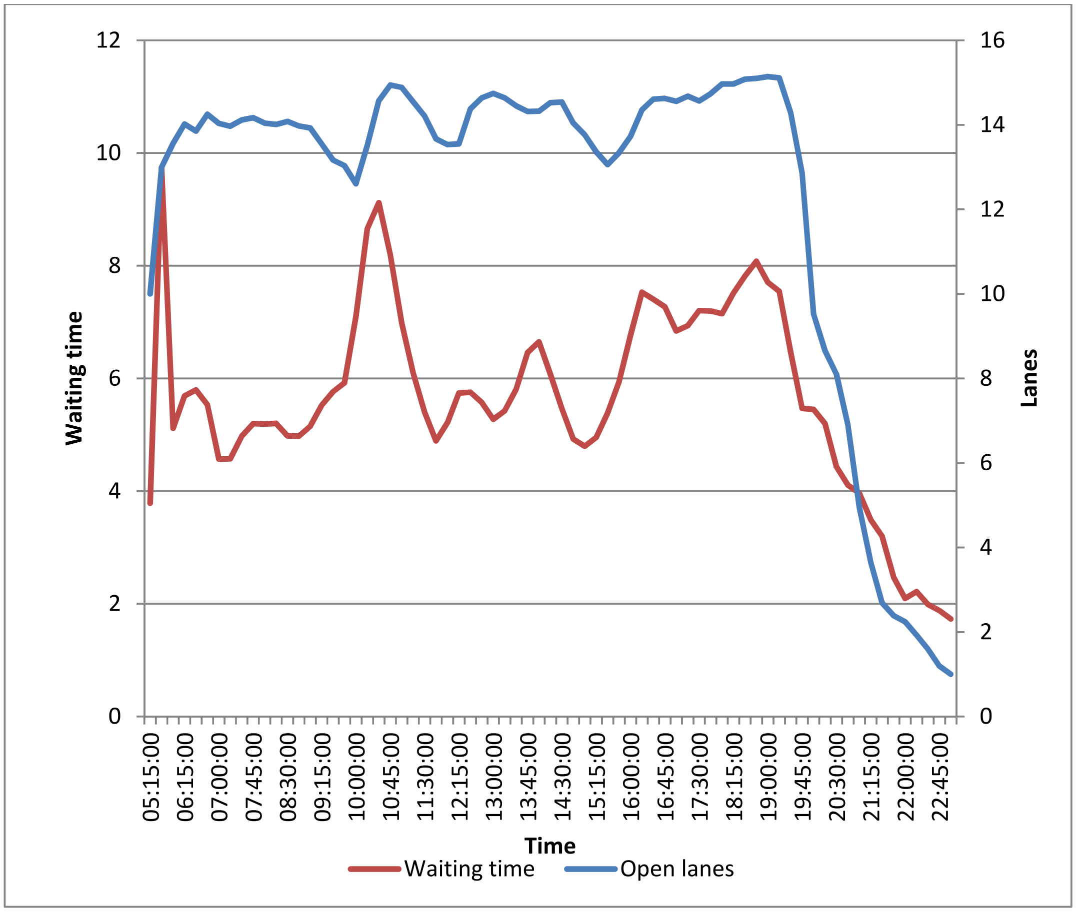

Figure 4 Average waiting time and average open lanes

Figure 5 shows the average waiting time as well as the 5% and 95% quantiles over the average day. 5% of waiting times are shorter than the 5% quantile, while 5% of waiting times are longer than the 95% quantile. It is seen that the variation is higher for 15 minute intervals where the average waiting time is high. But while the morning peak has the highest average waiting time, the longer waiting times as captured by the 95% quantile is seen to be higher in the midday peak and the afternoon peak. The 95% quantile is mostly 2 and sometimes 3 times larger than the average waiting time.

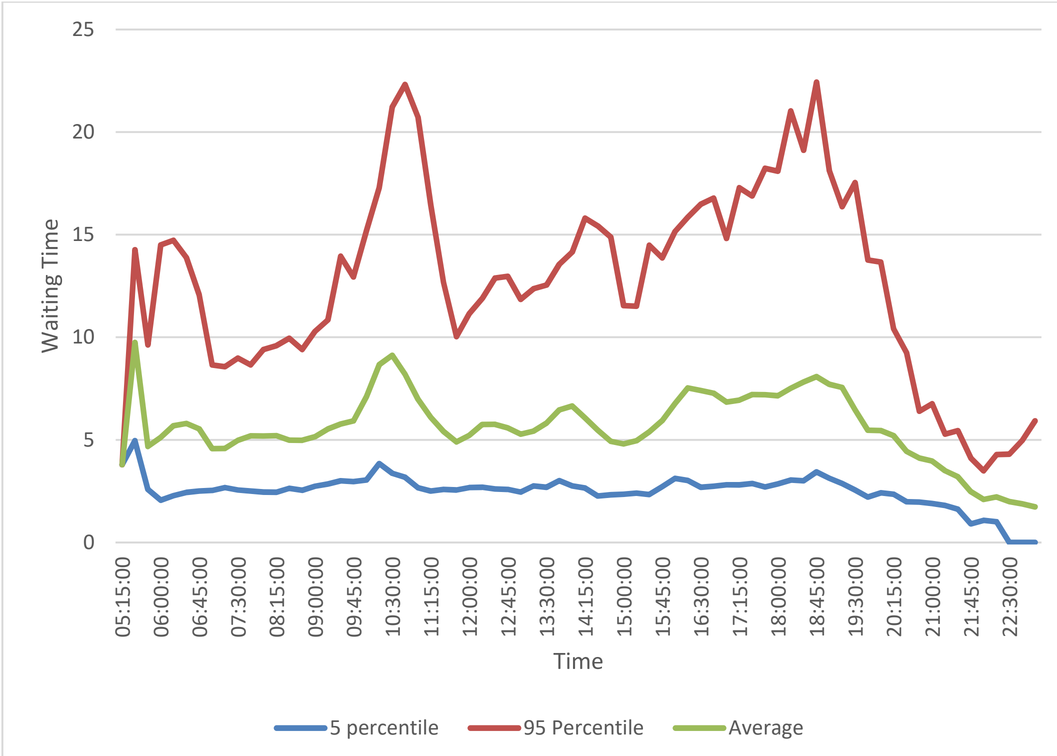

Figure 5 Average quarterly waiting time together with the 5% percentile and the 95% percentile across the day

Figure 6 shows the empirical cumulative distribution function of waiting time. The median waiting time is 4.20 minutes. 25% of waiting times are larger than 6.63 minutes, 10% of waiting times are larger than 10.80 minutes, while 5% of waiting times are larger than 14.11 minutes.

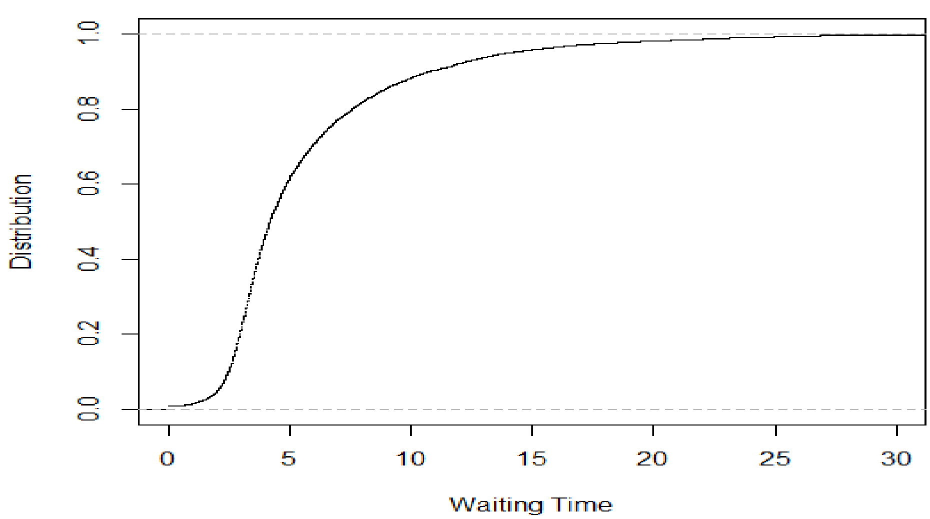

Figure 6 Empirical cumulative distribution of quarterly waiting times

3. Modelling

The modelling is based on the same models as described in report 1. Therefore this report focuses only on the things that are updated based on the new data, i.e. the reestimation, the validation of the new model, and the application of the new model in the evaluation.

To assess the effect on the waiting time cost of increasing capacity in the CSC, we need to estimate how the waiting time distribution depends on capacity. The analysis is performed using the same model as in phase 1. We refer to the phase 1 report for a description of that see Fosgerau et al. [1].

3.1 Estimation results

3.1.1 Mean regression

A number of regression models have been tested.

- First, we reestimated the same specification as used in report 1 on the new data. This gave similar results as before, which is reassuring.
- Then we split demand into three groups instead of two as we included low-cost carriers as a separate group.

The estimated regression parameters are shown in the following table. The dependent variable is log waiting time.

Table 1 Estimation results, mean regression

	Estimate	Std. Error	t value	
(Intercept)	0.49200	0.07330	6.71	***
log(\`Open lanes\`)	-0.07580	0.02210	-3.43	***
log(laglane)	-0.28600	0.02650	-10.79	***
log(laglane2)	0.20000	0.02290	8.73	***
OLoad_Sum2	0.00001	0.00003	0.39	
OLoad34	0.00002	0.00002	1.10	
OLoad56	0.00001	0.00002	0.53	
OLoad_7	0.00008	0.00003	2.69	**
OLoad_8	0.00009	0.00003	3.23	**
OLoad_9	0.00006	0.00003	2.18	*
OLoad_10	0.00008	0.00003	3.10	**
OLoad_11	0.00005	0.00003	1.99	*
OLoad_12	0.00008	0.00003	2.98	**
LLoad24	0.00001	0.00001	0.74	
LLoad_5	0.00013	0.00002	6.10	***
LLoad_6	0.00007	0.00002	2.68	**
LLoad_7	0.00008	0.00003	2.92	**
LLoad89	0.00000	0.00002	0.19	

ELoad25	0.00001	0.00001	0.52	
ELoad_6	0.00009	0.00002	4.74	***
ELoad_7	0.00007	0.00002	3.61	***
ELoad_8	0.00012	0.00002	6.69	***
ELoad_9	0.00009	0.00002	5.29	***
ELoad_10	0.00004	0.00002	2.67	**
log(lagwait)	0.80400	0.00694	115.85	***
day2	0.01270	0.01100	1.15	
day3	0.01300	0.01080	1.20	
day4	-0.00295	0.01060	-0.28	
day5	-0.01090	0.01070	-1.02	
day6	0.00196	0.01180	0.17	
day7	0.02280	0.01080	2.11	*
hour6	0.00703	0.06620	0.11	
hour7	-0.01900	0.06600	-0.29	
hour8	0.00957	0.06650	0.14	
hour9	0.05250	0.06660	0.79	
hour10	0.13700	0.06690	2.05	*
hour11	-0.03820	0.06720	-0.57	
hour12	0.03940	0.06720	0.59	
hour13	0.01940	0.06740	0.29	
hour14	-0.03580	0.06670	-0.54	
hour15	0.01160	0.06630	0.17	
hour16	0.05840	0.06620	0.88	
hour17	0.05180	0.06640	0.78	
hour18	0.06870	0.06650	1.03	
hour19	-0.05720	0.06610	-0.87	
hour20	-0.08730	0.06690	-1.30	
hour21	-0.22700	0.06900	-3.29	***
hour22	-0.31600	0.07070	-4.47	***
hour23	-0.57900	0.07800	-7.42	***

Multiple R-squared: 0.755. DW = 2.0008, p-value = 0.4202.

We make the following observations.

- "Open.lanes","laglane" and "laglane2" are the number of open lanes in the current period (at time t), 15 minutes before (t-1) and 30 minutes before (t-2). The parameters for the lagged variables are very significant.
- These parameters imply that an increase in the number of open lanes at time t reduces the waiting time at time t+1. The effect is reduced but still very significant at time t+2. In addition, there is the amplifying effect of the lagged waiting time.
- "LLoad" refers to departures that are served by low-cost carriers. "ELoad" and "OLoad" refer to destinations inside Europe and outside, respectively, that are not served by low-cost carriers. The numbers refer to forward lags: Thus "OLoad56" is the number of people leaving on planes departing during 5th and 6th 15 minutes time interval from the

current time, i.e. 75 to 105 minutes, after the current time, for destinations outside Europe.

- Demand for destinations outside Europe affects waiting times up to 3 hours prior to departure ("OLoad12"). Demand for destinations inside Europe affects waiting times up to 2 hours and 30 minutes prior to departure, and low-cost demand affects waiting times up to 2 hours and 15 minutes prior to departure. This makes sense as passengers go through security at varying times prior to their departure. Effects beyond these were small and therefore ignored in the model.
- We have included forward lags of demand until the point where the parameters become smaller and less significant.
- All the demand parameters are positive as expected.
- "lagwait" refers to the waiting time at time t-1. The parameter is 0.80, which means that a change in the waiting time, either due to random shocks or due to changes in the independent variables, will persist for some time into the future. The effect of a temporary change will die out over time, while the effect of a permanent change will take time to be fully reflected in the waiting time.
- The "day" constants take day-specific effects into account. Monday ("day1") is the base, and hence the constants measure the difference from Mondays. The differences between days, taking into account all the other variables in the model, are small and only some are significantly different from zero. The waiting time is longer on Sundays for reasons not otherwise explained in the model.
- The "hour" constants take time of day specific effects into account. The hour from 5am to 6am is the base and "hour" constants measure the differences from this hour. The effects are small, except from 9pm where waiting times decrease until the last hour from 11pm to 12pm. At 10 am the constant is also significant but smaller and positive.
- We have tested the residuals for auto-correlation. The Durbin-Watson test, as well as a range of other tests, do not reject the hypothesis of no serial correlation of the residuals.

3.1.2 Variance regression

We then estimated various specifications of the variance regression, arriving at the model with parameters shown in Table 2. Again, the model was extended until the point where the autocorrelation of the residuals could be assumed to be zero.

The dependent variable is the logarithm of the squared residuals from the mean regression. This measures the scale of the variability of waiting times. The unit for this does not have an easy interpretation.

Table 2 Estimation results, variance regression

	Estimate	Std. Error	t value	
(Intercept)	-2.78428	0.10304	-27.02	***
log(`Open lanes`)	-0.53547	0.03919	-13.66	***
hour23	1.11374	0.27035	4.12	***
laglogsquareresiduals_reg1	0.0455	0.01096	4.15	***

Multiple R-squared: 0.0368. DW =1.9991, p-value = 0.4799

The following observations can be made.

- Increasing the number of open lanes decreases the variability of waiting time. The effect is very significant.
- The variability is higher in the last hour before midnight.
- There is a tendency that a numerically large residual in one period (higher or lower waiting time than otherwise expected) is associated with increased variability also in the next period.
- The fit of the model, measured by the R-squared, is low. This is expected since the dependent variable is constructed from the random residual from the first-stage regression.
- The Durbin-Watson statistic, as well as a range of other tests, allows us to accept that residuals are not auto-correlated.

3.1.3 Model validation using simulation results

Simulation is carried out for an average week, covering the interval from 6am to 8pm. The profiles over the week of demand and the number of open lanes are constructed as the average over the weeks in the data. The construction of the demand variables proceeds as follows.

The data informs about the number of passengers according to the time of departure. These data are used as they are in the estimation and in the simulation of the model. For the purpose of evaluating the waiting time cost we need the number of passengers according to the time they pass through security. We distribute each departing passenger on earlier times using the demand coefficients from the estimated model. These coefficients are normalised to sum to 1, such that each departing passenger is counted exactly once at the central security check.

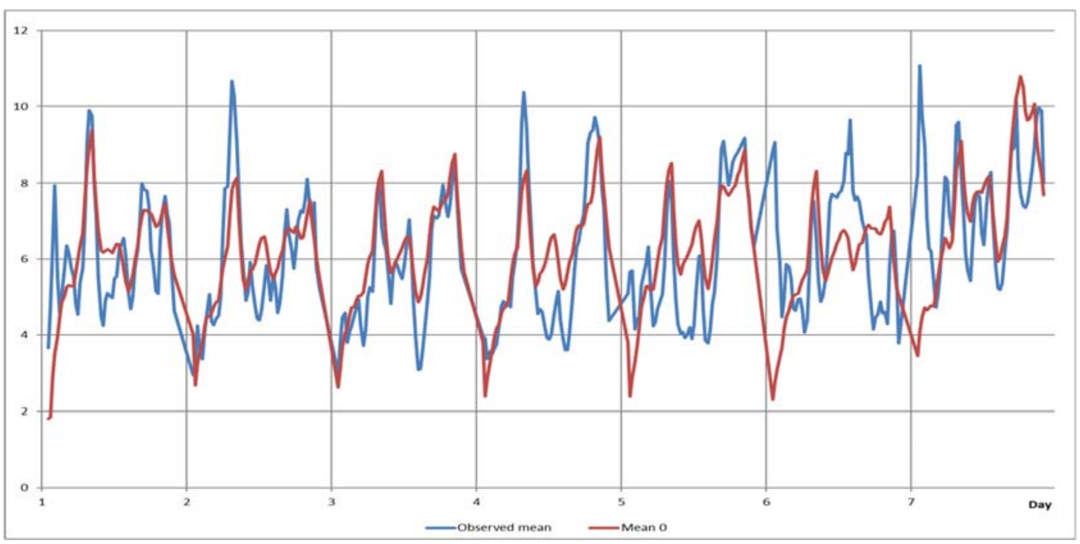

Figure 7 Comparison of predicted mean waiting time to the observed waiting time

Figure 7 compares the model prediction to the observed mean waiting times from 6 am to 8 pm. The model tracks the changes within days and over the week quite well. Deviations are to be expected, as there is sampling noise in the observed data. The model has a tendency to underpredict the mean waiting time. This is expected since the model is estimated in terms of log waiting time. The purpose of the model is to evaluate the change in waiting times following an increase in the number of lanes and then the bias does not matter much.

We obtain similar results when we compare the predicted and the observed standard deviation of waiting time, as seen in Figure 8. Larger differences between predicted and observed must be expected for the standard deviation of waiting times than for the mean, since sampling noise matters more for the standard deviation.

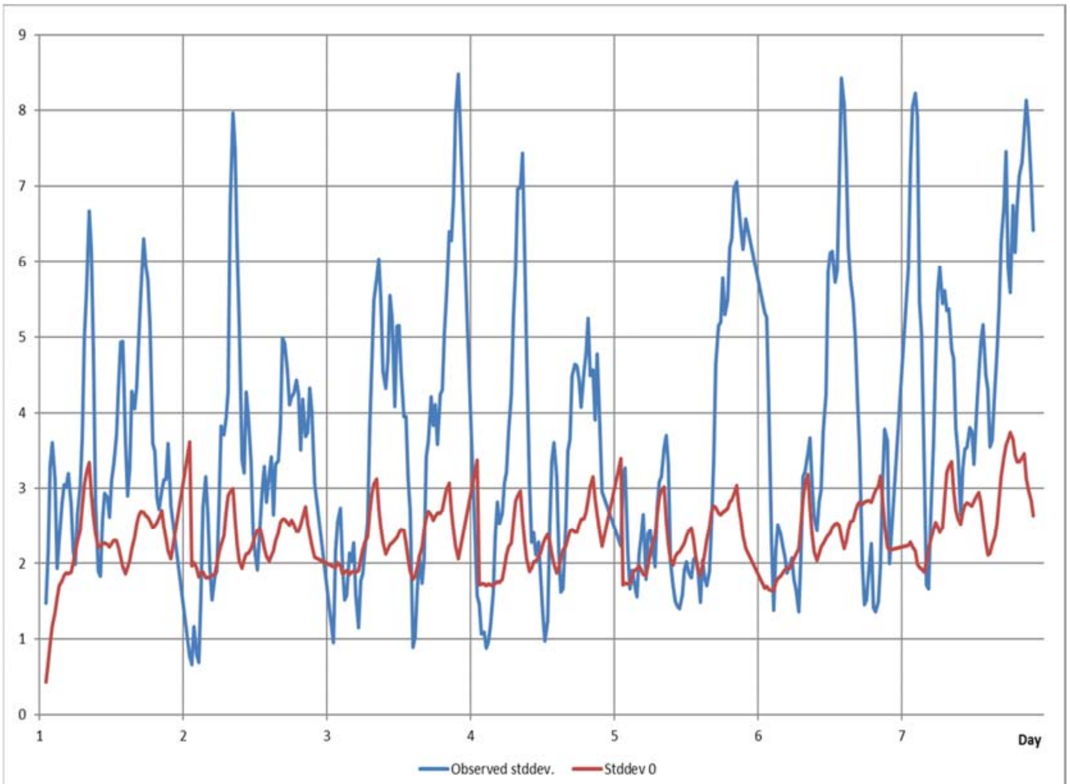

Figure 8 Observed and predicted standard deviation of waiting time

Finally, Figure 9 shows the 5% and the 95% quantile for the observed data and for the base scenario. This shows that the model is able to track the distribution of waiting times quite well. The match is not expected to be perfect, due to sampling noise in the observed data.

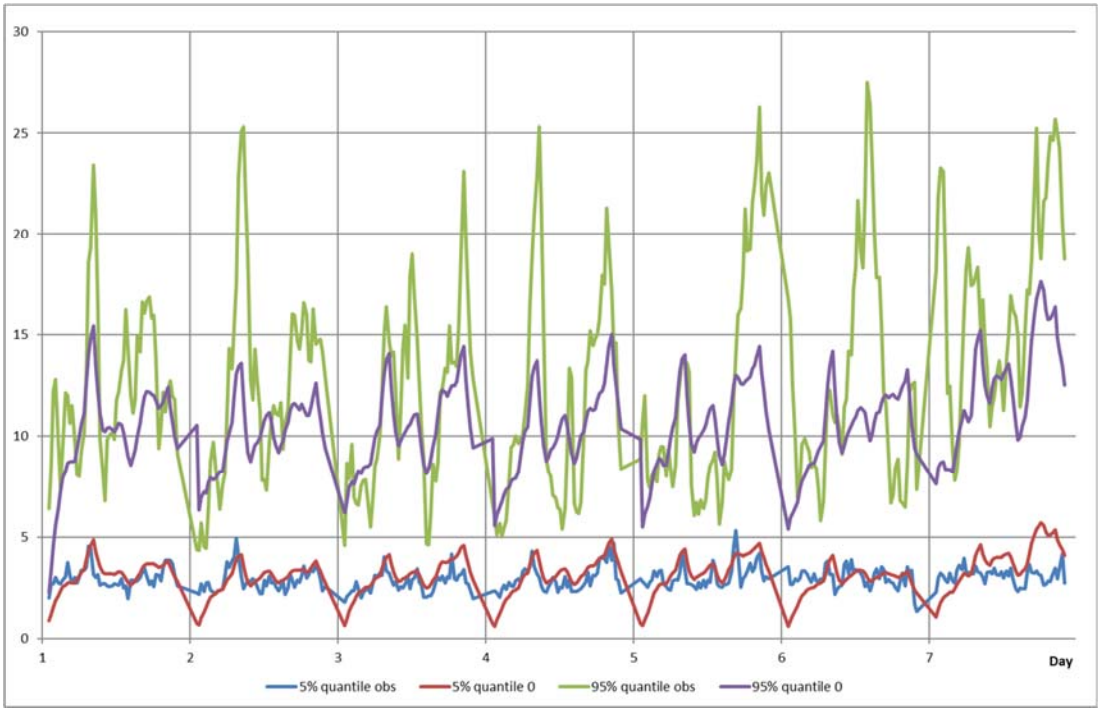

Figure 9 5% and 95% quantiles, observed and base scenario 0

In conclusion, we find that the model matches the observed data quite well: it is able to track the mean, the standard deviation and the range of the distribution of waiting time quite well over the simulated week. This is very satisfactory.

4. Simulation results

We simulate four scenarios, using 50,000 replications, for an average week and from 6am to 8pm every day. We omit periods outside this interval as the number of passengers and the number of open lanes are low; thereby we ensure that the simulation stays within the range of data where the model is most reliable.

The scenarios are based on the demand profile and the profile for the number of open lanes in the central security check. The base scenarie 0 is the average number of lanes across the weeks in the data. Scenarios 1, 2, 3 add 1, 2, and 3 additional lanes, respectively. A cap of 20 lanes is applied, which is the maximum available at the airport.

The following table shows the mean and the standard deviation of waiting time for the four scenarios as well as for the observed data. Adding one lane to the base scenario at all times decreases the mean waiting time by 0.42 minutes. Adding more lanes decreases the mean waiting time further at a diminishing rate, which is as expected and reasonable. Similarly for the standard deviation of waiting time, the first lane added leads to a decrease of 0.23 minutes and more lanes decrease the standard deviation of waiting time further.

Table 3 Mean and standard deviation of waiting time 6am-8pm, minutes

Scenario	Mean	Standard deviation
Observed	6.17	3.61
Sim 0	6.28	2.36
Sim 1	5.86	2.13
Sim 2	5.51	1.98
Sim 3	5.22	1.81

5. Economic evaluation of a capacity increase

We use the same three coefficients as in report 1

$$\alpha = 3\ DKK/min, \beta = 4.5\ DKK/min, and\ \gamma = 13.5\ DKK/min$$

These are conservative estimates, since they omit the cost associated with missing a flight.

5.1 The cost of capacity

The cost of adding an additional lane for 1 hour with 4 employees has been calculated by the Danish Transport, Construction and Housing Agency [2] to be DKK 274 *4 = DKK 1096 with an uncertainty of +-10%.

The cost-benefit analysis is carried out in market prices, which corresponds to the willingness-to-pay of leisure travellers. The cost to the airport is in factor prices and must be converted to market prices. We use the standard factor of 1.325, which yields a cost of an additional lane in market prices of 1452 DKK per hour.

5.2 Comparing costs and benefits of a capacity increase

Figure 10 shows the waiting time cost per passenger in DKK for the four scenarios. We observe a consistent decrease over the week as more lanes are added.

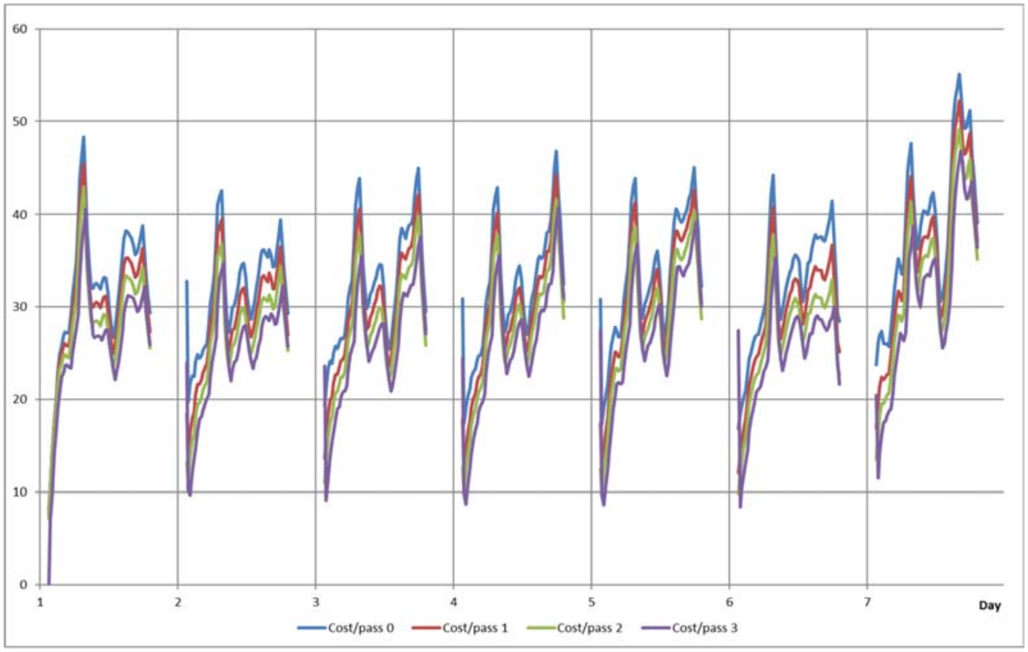

Figure 10 Waiting time cost per passenger, DKK

The average reduction in waiting time cost per passenger in the four scenarios is shown in the following table.

Table 4 Waiting time cost per passenger, DKK

Scenario	Cost per passenger, DKK	Change from previous
Sim 0	33.42	
Sim 1	30.87	2.55
Sim 2	28.82	2.05
Sim 3	27.10	1.72

The cost per passenger is multiplied by the number of passengers passing security every 15 minutes, estimated from flight departure data as explained above. The following figure shows the total savings from adding lanes in the central security check, comparing scenarios with additional lanes to the base scenario 0.

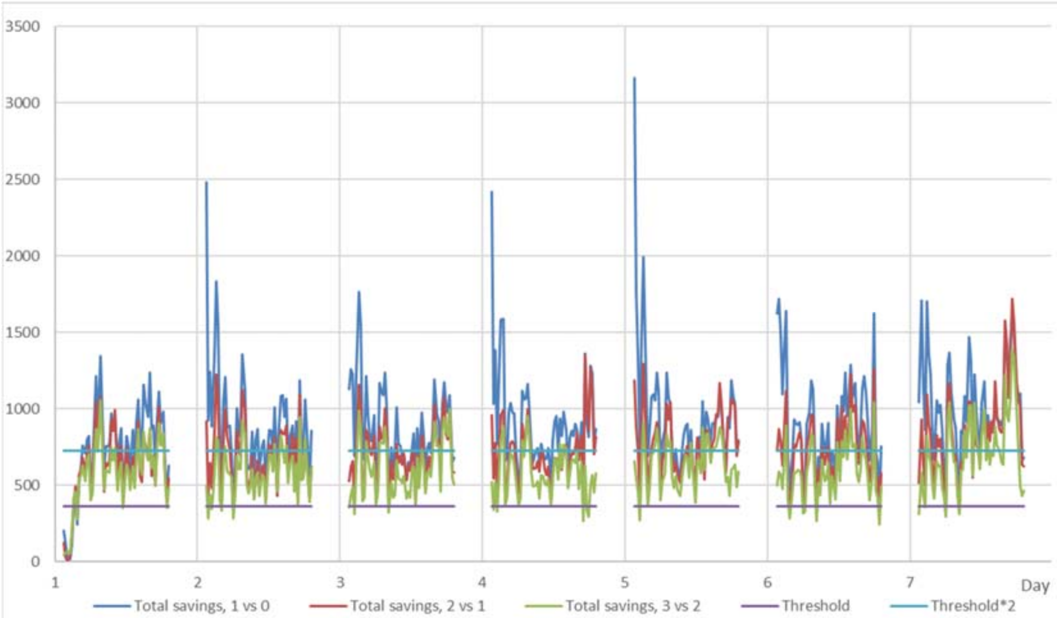

Figure 11 Total cost savings for passengers over an average week. The figure shows the benefit to passengers in 15 minute intervals of adding one lane. "1 vs 0" compares one additional lane to the baseline, "2 vs 1" compares two additional lanes to one additional lane, and "3 vs 2" compares three additional lanes to two additional lanes. The threshold lines indicate the corresponding cost to the airport per 15 minutes and twice that amount.

Adding 1 or 2 lanes yields a net benefit at all times during the week. Adding three lanes yields a substantial net benefit in general, but there are a few 15 minute intervals during the week where the net benefit of the third lane becomes negative.

Table 5 Benefits and costs per hour due to an additional lane open at all times during the day

	Savings to passengers	Savings due to mean waiting time	Cost to airport	Net benefit
From 0 to 1 extra lanes	4947	2383	1452	3495
From 1 to 2 extra lanes	3981	2001	1452	2529
From 2 to 3 extra lane	3338	1684	1452	1886

Opening one additional lane yields an average benefit to passengers of 4947 DKK per hour with a corresponding cost to the airport of 1452 DKK per hour. A bit less than half of the savings to passengers, 2383 DKK, is due to reduction in the mean waiting time, the rest is due to reduced variability. This means that adding one lane is justified even without taking the reduction in variability into account.

Opening a second additional lane yields an additional average benefit to passengers of 3981 DKK per hour, which is still larger than the cost to the airport of 1452 DKK per hour of manning a second additional lane. The saving due to mean waiting time is 2001 DKK, is also higher compared to the hourly cost for manning a lane.

Opening a third additional lane yields an additional average benefit to passengers of 3338 DKK per hour, out of which 1684 DKK per hour is due to reduction of the mean waiting time. The benefit to passengers clearly outweighs the cost to the airport of 1452 DKK per hour, also in this case going from two to three additional lanes.

In conclusion, the calculations indicate a clear net benefit of opening three additional lanes at all times during the day. The main points relevant for assessing the robustness of the conclusion are the following.

The benefits are proportional to the values of time and reliability. The values used are on the low side of the available evidence, which indicates that the actual benefit of a capacity increase is likely to be larger than the calculated benefit.

The cost per lane hour is directly proportional to the estimate from the Danish Transport, Construction and Housing Agency. The cost per lane hour would thus need to be more than 100% larger in order to change the conclusion.

We therefore find that the conclusion that there is a clear net benefit of opening three additional lanes at all times during the day is quite robust.

6. Conclusion

The results of phase 3 confirms the results found on earlier data in phase 1 that increasing capacity in the security check is beneficial to society even when costs are taken into account. This conclusion seems to hold for at least three additional lanes where the reduction in mean waiting time and waiting time variability more than outweight the cost of increased capacity.

References

[1] M. Fosgerau, S. L. Mabit, and A. Ranjan, “Waiting times in Copenhagen Airport - An economic evaluation of delays in the central security check,” Transport DTU, 2016.

[2] Trafik- og Byggestyrelsen, “Beregning af omkostninger ved at åbne et ekstra spor i den ”Centraliserede sikkerhedskontrol ”(CSC) i Københavns Lufthavn,” Journal TS31200-00182, Dec. 2015.