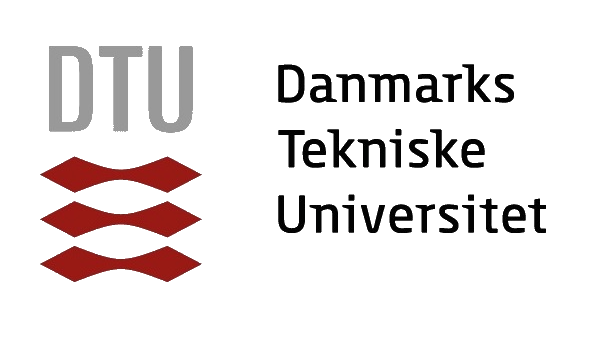
**42178**

Transport system analysis – demand and planning

Portfolio 3



**s212156 -** Irene Gómez Sánchez

**s212904 -** Zoltán György Varga

**s213470 -** Jakob Ryttergaard Poulsen

Content

[Table list 2](#_Toc87362893)

[Exercise 1 3](#_Toc87362894)

[Section 1: Estimation 3](#_Toc87362895)

[Section 2: Market shares 4](#_Toc87362896)

[Section 3: Value of Time 4](#_Toc87362897)

[Section 4: Elasticities 5](#_Toc87362898)

[Section 5: Model specification 5](#_Toc87362899)

[Section 6: Simulation, scenario 1 6](#_Toc87362900)

[Section 7: Simulation, scenario 2 7](#_Toc87362901)

[Exercise 2 8](#_Toc87362902)

[Section 1: Model discussion and selection 8](#_Toc87362903)

[Section 2: Model structure 9](#_Toc87362904)

[Section 3: Cost elasticities 10](#_Toc87362905)

[Section 4: Scenario and calibration 11](#_Toc87362906)

[References 13](#_Toc87362907)

[Appendixes 13](#_Toc87362908)

[Appendix A 13](#_Toc87362909)

# Table list

[Table 1. Exercise 1.1 AIC results 3](#_Toc87362910)

[Table 2. Exercise 1.1 Rho squared 4](#_Toc87362911)

[Table 3. Exercise 1.2. Market shares 4](#_Toc87362912)

[Table 4. Exercise 1.2. Expected vs observed mode choices 4](#_Toc87362913)

[Table 5. Exercise 1.3. VOT 5](#_Toc87362914)

[Table 6. Exercise 1.4. Cost and time elasticity matrixes 5](#_Toc87362915)

[Table 7. Exersice 1.5 Market share 6](#_Toc87362916)

[Table 8. Exercise 1.6. Market shares 6](#_Toc87362917)

[Table 9. Exercise 1.7. Market shares 7](#_Toc87362918)

[Table 10. Exercise 2.1. AIC Values 9](#_Toc87362919)

[Table 11. Exercise 2.1. Within-nest conditional probabilites 9](#_Toc87362920)

[Table 12. Exercise 2.1. Probability of choosing a nest 10](#_Toc87362921)

[Table 13. Exercise 2.1 Overall probabilities 10](#_Toc87362922)

[Table 14. Exercise 2.1. Market shares 10](#_Toc87362923)

[Table 15. Exercise 2.1. VOT 10](#_Toc87362924)

[Table 16. Exercise 2.3. Cost and time elasticity matrixes 11](#_Toc87362925)

[Table 17. Exercise 2.4. Market shares 12](#_Toc87362926)

[Table 18. Exercise 2.4 ASCs 12](#_Toc87362927)

[Table 19. Exercise 2.4. Final hypothetical scenario market shares 12](#_Toc87362928)

[Table 20. Exercise 2.4. Results from ASC iterative procedure 13](#_Toc87362929)

# Exercise 1

## Section 1: Estimation

**The following two logit models based on utility functions 1 and 2, respectively, have been implemented in R.**

Texto, Carta

Descripción generada automáticamenteTexto

Descripción generada automáticamente

**Discuss the results and decide which of the specifications you prefer. Argue why. The remainder of the exercise, you should only base on your preferred specification.**

There are two utility functions which reflect different models trying to explain the same thing based on the same data. None of them is a restricted nor extended version of the other one, because we cannot get one model from the other by setting coefficients to 0, due to the log transformation of TC in utility function 2.

The discussion of the results is as follows:

For Vn1: From the constants we can note that, excluding time and cost, and taking walk as the reference, the utility decreases when using the bike, increases when using the car and decreases when using public transport. Moreover, regarding ***a*** parameter, its negative value means that increasing the cost of the trip would decrease the utility, and the same with ***b***. Both signs make sense, however, we cannot compare their sizes since they have different units.

For Vn2: again, excluding time and cost, and taking walk as the reference, the utility increases when using the bike, increases when using the car and decreases when using the public transport. Both signs of ***a*** and ***b*** being negative again make sense, however, we still cannot compare their sizes since they have different units.

Next, the selection of the preferred model:

Since one model is not a restricted version of the other, we cannot use Log-Likelihood ratio. However, we can calculate the AIC value and get the lower one. AIC formula is:

Then, we get the following results:

|  |  |
| --- | --- |
| AIC model 1 | 8552.2 |
| AIC model 2 | 8524.8 |

Table 1. Exercise 1.1 AIC results

Model 2 has a smaller value of AIC than model 1, therefore, based on this criterion, we would choose the model 2.

Besides, we could calculate rho squared for comparing both models, which value is 0.003 according to Table 2. This small value indicates that the model 2 is better than model 1.

|  |  |
| --- | --- |
| ρ2 | 0.003 |

Table 2. Exercise 1.1 Rho squared

## Section 2: Market shares

**Based on the estimated parameters, calculate the individual probabilities for each alternative in each observation. Compute the market shares for each mode in the base situation and compare to the observed market shares.**

As asked, we created the utilities needed to calculate the probabilities. From the data given in the heading, we see that K1=0 for identification. Then, utility functions are obtained as follows:

And, afterwards, individual choice probabilities are calculated by the following formula:

For each individual we get four probabilities, since there are four mode choices. The average probability for each mode is shown in the Table 3. These percentages are the market shares, and they should sum to 1.

|  |  |  |  |
| --- | --- | --- | --- |
| Pav(1) | Pav(2) | Pav(4) | Pav(6) |
| 0,194 | 0,223 | 0,547 | 0,036 |

Table 3. Exercise 1.2. Market shares

In the following table, Table 4, expected and observed market shares are shown. We see that both are the same, which is because the logit model has been estimated based on the observed data. So it is possible to accurately approximate utility functions that describes a data-set.

Chart, bar chart

Description automatically generated

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | MS1 | MS2 | MS4 | MS6 |
| Observed | 19.4% | 22.3% | 54.7% | 3.6% |
| Expected | 19.4% | 22.3% | 54.7% | 3.6% |

Table 4. Exercise 1.2. Expected vs observed mode choices

## Section 3: Value of Time

**In mode choice models, we can calculate the value-of-time (VOT). This is equal to:**

Diagrama, Pizarra

Descripción generada automáticamente

**where 𝑡 and 𝑐 refer to travel time and travel cost. Calculate the average VOT for the sample in the data.**

Since there are four modes, we get four VOT. According to the formula, the VOT for each mode is obtained as follows:

The average VOT for each mode is shown in the Table 5.

|  |  |  |  |
| --- | --- | --- | --- |
| VOT1 | VOT2 | VOT4 | VOT6 |
| 0.064 | 0.129 | 0.512 | 0.844 |

Table 5. Exercise 1.3. VOT

The units are in DKK/min.

When we calculate the value of time like this, it is most correct to weight them by the probabilities of people choosing this alternative. So for each observation we take the estimated VOTi and multiply it with the probability of this person choosing alternative i. Then we divide it with the sum of probabilities.

A picture containing text, person, watch, gauge

Description automatically generated

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 0.0237 | 0.0733 | 0.6805 | 0.8740 |

We see, as expected that the highest value of time is for public transport, then cars, then bikes and lastly walking.

## Section 4: Elasticities

**For travel cost and time, calculate the aggregate direct and cross elasticities for each mode based on probability-weighted averages. Present these aggregated results, e.g. in a four by four elasticity matrix and discuss the results.**

To calculate aggregated elasticities, we cannot obtain them as the average of the elasticities, but we must weight them with the probability of each observation choosing their mode choice. Table 6 shows the cost and time elasticity matrix.

For a logit model, we can calculate direct elasticities, for a log-linear model as:

And cross-elasticities as:

When the model relationship is not log-linear, we need to consider the table below:

Table

Description automatically generated

So we get that

And

We calculate the elasticities for all modes, for all observations. And we take the probability-weighted average of all the elasticities, using the formula below:

Text, letter

Description automatically generated

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cost elasticity matrix | | | | |
|  | 1 | 2 | 3 | 4 |
| 1 | -0,33211 | 0,1108207 | 0,2685499 | 0,0212929 |
| 2 | 0,1168214 | -0,338111 | 0,2685499 | 0,0212929 |
| 3 | 0,1168214 | 0,1108207 | -0,180382 | 0,0212929 |
| 4 | 0,1168214 | 0,1108207 | 0,2685499 | -0,427639 |
| Time elasticity matrix | | | | |
|  | 1 | 2 | 3 | 4 |
| 1 | -0,438839 | 0,0739117 | 0,1726471 | 0,0331107 |
| 2 | 0,0944462 | -0,313868 | 0,1726471 | 0,0331107 |
| 3 | 0,0944462 | 0,0739117 | -0,065462 | 0,0331107 |
| 4 | 0,0944462 | 0,0739117 | 0,1726471 | -0,565496 |

Table 6. Exercise 1.4. Cost and time elasticity matrixes

We can note that the diagonals of both matrixes are made of negative numbers, which make sense because they are direct elasticities. Therefore, if the time or the cost of any of the modes increased, its specific demand would decrease.

Nevertheless, the rest of the elasticities in the matrixes, which are the cross-elasticities, are positive values. This also makes sense since they mean that an increase of time or cost in any of the modes would lead to an increase in the demand of the other modes.

Although the sum of the elasticities does not sum to 1, the sum of the net variations in demand must sum to 1, since the total demand must remain constant.

## Section 5: Model specification

**Now, you would like to investigate whether there is a difference in preferences for using public transport between male and female respondents. Make a cross table presenting how the market shares are split across gender. In the output presented below, the interaction variable Ptfem has been included, which is only one in the public transport alternative and only if the user is female:**

Imagen de la pantalla de un celular con texto

Descripción generada automáticamente con confianza media

**Discuss the results. Are the market shares predicted for the new model different than those predicted by the model in exercise 1.2? Why/Why not?**

We calculated the market share using the below presented equitation. To create the gender variable, we subtracted 1 from every value of the original gender variable.

The calculated market share, the observed market share and the market share which was calculated during exercise 1.2 are presented in Table 7.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | MS1 | MS2 | MS4 | MS6 |
| Observed | 19.4% | 22.3% | 54.7% | 3.6% |
| Expected Exercise 1.2 | 19.4% | 22.3% | 54.7% | 3.6% |
| Expected Exercise 1.5 | 19.4% | 22.3% | 54.7% | 3.6% |

Table 7. Exersice 1.5 Market share

The market share is the same what we have predicted during exercise 1.2. The reason of this, the gender has a little influence at the market share.

So the overall market shares remain the same. How about the gender distributions?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Female-percentages** | MS-walk | MS-bike | MS-car | MS-public |
| Observed | 67.4% | 56.3% | 46.7% | 72.8% |
| Expected Exercise 1.2 | 57.4% | 55.7% | 51.9% | 51.1% |
| Expected Exercise 1.5 | 56.7% | 55.1% | 51% | 72.8% |

So, while the overall market-changes have not been affected by the inclusion of a female public-transport coefficient, the public transport market shares have more accurately approximated the distribution of genders that we can observe in the dataset.

## Section 6: Simulation, scenario 1

**Now, politicians would like you to evaluate what happens with market shares if the cost of going by car increases by as much as 50%! Calculate the new market shares and comment on your results; ①.**

If the cost of going by car increased by 50%, the input of the model should change so TC4 is now 1.5 times the original one. In this case, Vn4 changes as well and, therefore, probabilities are different, which leads to a different market shares.

The new market shares are as in Table 8.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | MS1 | MS2 | MS4 | MS6 |
| Original | 19.4% | 22.3% | 54.7% | 3.6% |
| Sim1: TCcar | 21.0% | 24.3% | 50.7% | 4% |
| ***Difference*** | ***+1.6%*** | ***+2%*** | ***-4%*** | ***+0.4%*** |

Table 8. Exercise 1.6. Market shares

We note a decrease in car market share, which seems logical since the cost has increased and the coefficient of ln(cost) in the model is negative. This means that an increase in the cost leads to a decrease in the utility, and that is the reason why less people use it.

We also observe that the other three mode choices increase their market shares.

## Section 7: Simulation, scenario 2

**A new more environmentally friendly car enters the world market. Unfortunately, it is more expensive and slower than the present car. Given that you have already implemented the car cost increase in 6) (for ordinary cars), it is still 20% more expensive than the ordinary car and 20% slower. What will be the market share for the new automobile according to your model? ②**

**Describe how the new alternative mode “steals” market shares from the other modes. What is the relative change in market share for each alternative? Is this reasonable given that the new mode is quite similar to the car mode? What would be a reasonable substitution pattern?**

In this case, the scenario sets out to introduce a new mode choice (new car), which is going to be included in the model as a new variable, X7. As Hint 2 indicates, the parameters of X7 are the same as X4. Besides, its new characteristics should be included: it is 20% more expensive, so TC7 is 1.2 times TC4 in section 6 (which was 1.5 times the original one); and it is also 20% slower, so TT7 is 1.2 times TT4.

Since there is a new mode, a new utility function is defined:

And therefore, now exists the probability of people choosing this mode, which changes the market shares as shown in the Table 9.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | MS1 | MS2 | MS4 | MS6 | MS7 |
| Original | 19.4% | 22.3% | 54.7% | 3.6% |  |
| Sim2: TC7 | 15.4% | 17.4% | 34.3% | 2.7% | 30.2% |
| ***Difference*** | ***-4%*** | ***-4.9%*** | ***-20.4%*** | ***-0.9%*** | ***+30.2%*** |

Table 9. Exercise 1.7. Market shares

In this case, we note that the new car takes a great market share (30.2%) at the expense of all the other mode choices, whose market shares decrease because the total amount of people travelling is now distributed into more choices. In greater detail, we observe that the market share of conventional car drivers decreases more than the other cases. This could be explained by the fact that both are cars, so they share similarities and attract people with similar profiles (Rich J. & Mabit, 2018).

# Exercise 2

**In order to test if there are correlation among some pairs of alternatives, the following nested logit models have been estimated:**

**Individual vs. Public (1,2,4 vs. 6)**

**Motor vs. other (1,2 vs. 4,6)**

**Car vs. others (4 vs. 1,2,6)**

**Car and Walk vs. others (1,4 vs 2,6)**

## Section 1: Model discussion and selection

**Based on the following estimation output, discuss the results and decide which specification you prefer:**

*Text

Description automatically generatedText

Description automatically generated*

Text

Description automatically generatedText

Description automatically generated  
Now we are looking at nested logit models. Nested logit models are very similar to logit models, except for an additional term in the model, the nest coefficients. As with multinomial logit models (MNL), the principle behind nested logit models is random utility maximisation (RUM). We construct utility functions as in the MNL

Where is the systematic utility and are error terms. For a nested logit model, we extend the utility function, by splitting the systematic utility into systematic utility shared across options within a nest and , alternative-specific systematic utility for option within the nest. For our given models, becomes 0, as there are no shared utility between the options within the nests.

The probability of choosing an option is now the product of two probabilities, the probability of choosing an option, conditional on choosing the nest, and the general probability of choosing the nest.

And the conditional probability of choosing option within a nest is given as:

The general probability of choosing nest is given by

Text

Description automatically generated with medium confidence

And the terms are defined as the nest-specific logsums.

A picture containing text

Description automatically generated

And the final probability of choosing option can now be written as

A picture containing text

Description automatically generated

If we look at the models above, we see that all of the signs for cost, time and PTfem are the same; negative signs for time and cost, as we expect and an increase in utility for females taking public transport, as explored in the section above. All effects are also statistically significant.

Regarding the nest coefficients, we can immediately see that from the equation above. Furthermore, if the nest coefficient is 1, the model has maximum substitution and converges as a multinomial logit model with no nesting structure. For nest coefficients above 1, the model is also not valid, so we add that .

From looking at the nesting coefficients criteria we can immediately exclude model 2, as the nesting coefficients are above 1. Model 1 and model 3 is structurally the same model, but with different estimations of parameters, so here we can choose by the log likelihood value closest to zero. Model 4 has a different nesting structure, and we can thus not say that model 4 and model 1 or 2 is a restricted version of the other, which disqualifies log-likelihood ratio tests. We can instead use the Akaike Information Criteria (AIC) as we did before and choose by the lowest value.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Model 1 | Model 3 | Model 4 |
| AIC | 8495.8 | 8497.0 | 8490.8 |

Table 10. Exercise 2.1. AIC Values

And we choose model 4 by the lowest AIC value. This is a nesting structure that splits *walking*(mode 1)and *car*(mode4) into nest 1 and *bike*(mode2)and *public transport*(mode6) into nest 2.

## Section 2: Model structure

**Based on the estimated parameters for the preferred model, calculate the individual probabilities for each alternative in each observation. Compute the aggregate market shares and VoT for each mode in the base situation.**

First we calculate the within-nest conditional probabilities by the following formula

And we find

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 0.27 | 0.73 | 0.84 | 0.16 |

Table 11. Exercise 2.1. Within-nest conditional probabilites

Then we calculate the probability of choosing a nest by

Text

Description automatically generated with medium confidence

Note that . And we find

|  |  |
| --- | --- |
|  |  |
| 0.74 | 0.26 |

Table 12. Exercise 2.1. Probability of choosing a nest

The overall probabilities are now found and shown in Table 13.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 0.19 | 0.22 | 0.55 | 0.04 |

Table 13. Exercise 2.1 Overall probabilities

The market shares are the percentages of the previous values and are shown in Table 14.

|  |  |  |  |
| --- | --- | --- | --- |
| MS1 | MS2 | MS4 | MS6 |
| 19% | 22% | 55% | 4% |

Table 14. Exercise 2.1. Market shares

The value of time is defined as

And the resulting unit is in money/time (e.g., dkk/hour)

The value of time equation for mode in our model becomes

And are found to be

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| -0.048 | -0.003 | 0.105 | 0.172 |

Table 15. Exercise 2.1. VOT

We see that it follows the same order of magnitudes as seen previously; from lowest to highest: walking, biking, car, public transport.

We can also take the proability weighted averages

A picture containing text, person, watch, gauge

Description automatically generated

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| -0.103 | -0.028 | 0.123 | 0.175 |

## Section 3: Cost elasticities

**Now, politicians would like to know the cost elasticities related to car cost and time (both direct and cross). You can derive the elasticities based on a simulation where you increase car cost (time) by either 1% or 10%. Compare these to the elasticities from exercise 1.4.**

To calculate the elasticity by simulation, first we need to take the market shares for a mode of transport , then we increase an underlying variable by 10%, and reevaluate the market share . These market shares are evaluated using the nested logit model above by multiplying the market-shares with .

Where

If the variable we increase is car cost, then the equation gives direct elasticity for car market share, with respect to cost, and we expect it to be negative. For other modes, it gives you the cross-elasticity with respect to car cost. Using this method we get the following elasticities:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | walk | bike | car | pub |
| Elasticity(carcost) | 0.2049644 | 0.1503933 | -0.1463257 | 0.1906525 |

As we expect, the cross elasticity is positive, but the direct elasticity for car cost to car market share is negative. When car cost increases, we expect the market share of cars to decline, and the other market shares to increase.

And for time:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | walk | bike | car | pub |
| Elasticity(cartime) | 0.06475367 | 0.0631028 | -0.05962632 | 0.1678881 |

Again, the signs are as we expect, negative for cars and positive for all other options.

The calculated elasticities are presented in Table 16. To calculate aggregated elasticities, we cannot obtain them as the average of the elasticities, but we must weight them with the probability of each observation choosing their mode choice

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cost elasticity matrix | | | | |
|  | walk | bike | car | Public transport |
| Walk | -0.2323 | 0.0770 | 0.1875 | 0.0194 |
| bike | 0.0831 | -0.2383 | 0.1875 | 0.0194 |
| car | 0.0831 | 0.0770 | -0.1279 | 0.0194 |
| Public transport | 0.0831 | 0.0770 | 0.1875 | -0.2960 |
| Time elasticity matrix | | | | |
|  | walk | bike | car | Public transport |
| walk | -0.3298 | 0.0588 | 0.1294 | 0.0378 |
| bike | 0.0713 | -0.2477 | 0.1294 | 0.0378 |
| car | 0.0713 | 0.0588 | -0.0517 | 0.0378 |
| Public transport | 0.0713 | 0.0588 | 0.1294 | -0.4638 |

Table 16. Exercise 2.3. Cost and time elasticity matrixes

Again, we can note that the diagonals of both matrixes are made of negative numbers, which make sense because they are direct elasticities. Therefore, if the time or the cost of any of the modes increased, its specific demand would decrease.

Nevertheless, the rest of the elasticities in the matrixes, which are the cross ones, are positive values. This also makes sense since they mean that an increase of time or cost in any of the modes would lead to an increase in the demand of the other modes.

Although the sum of the elasticities does not sum to 1, the sum of the net variations in demand must sum to 1, since the total demand must remain constant.

Comparing the elasticities, the cost elasticities become closer to 0 compare the elasticities in exercise 1.4 and the Time elasticities become closer to 1.

## Section 4: Scenario and calibration

**Again, you are asked to analyse what happens when a new more environmentally friendly car enters the market. It is more expensive and slower than the present cars. It is 20% more expensive than the ordinary cars and 20% slower. Discuss how to include the new alternative in your model! You are also informed that aggregate population shares for each mode are:**

**walk - 0.15, bike - 0.25, car - 0.39, pub - 0.08, and new car - 0.13.**

**Please update your model so that it includes the new alternative and meets the aggregate shares on the first two decimals of the shares. What are the new ASCs in your updated model? ③**

**Discuss the substitution patterns of you updated model. What will happen in a scenario where the new\_car reaches the same travel cost and travel time as the existing car? You are welcome to illustrate this either by calculating market shares in the scenario or calculate elasticities of the updated model.**

**Hint ③: In this scenario, assume that the preferences (cost and time parameters) for the new alternative are the same as for car. You can then calibrate the alternative specific constants using algorithm 15.1.**

In this case, again, the scenario sets out to introduce a new mode choice (new car), which is going to be included in the model as a new variable, X7. Since it is similar to cars, it would enter as a new alternative in the Walk and Car nest.

As Hint 3 indicates, the preferences of X7 are the same as X4. Besides, its new characteristics should be included: it is 20% more expensive, so TC7 is 1.2 times TC4; and it is also 20% slower, so TT7 is 1.2 times TT4.

Since there is a new mode, a new utility is defined:

We use an iterative calibration algorithm of the form

Where is the observed market shares we wish to calibrate for, and is the market shares calculated by the model using . When has approximated the algorithm converges.

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated Graphical user interface, text, application, email

Description automatically generated Graphical user interface, text, application, email

Description automatically generated Graphical user interface, text, application, email

Description automatically generated Graphical user interface, text, application, email

Description automatically generated

If we use the calibrated model and set cost and time for the new car equal to the cost and time for the old car, we get the following market shares:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | MS-walk | MS-bike | MS-car | MS-pub | MS-newcar |
| Given | 15% | 25% | 39% | 8% | 13% |
| Estimated\_scenario | 14.8% | 24.7% | 38.1% | 7.8% | 14.4% |
| Difference | -0.2% | -0.3% | -0.9% | -0.2% | +1.4% |

# References

Rich J. & Mabit, S. E., 2018. *Transport Models - From Theory to Practise.* 9 ed. Lyngby, Denmark: s.n.