Markov Chains Evolution and The Role of Dynamic Programming in Multi-way Choices: An Analysis of SARS-CoV-2 Data

by

N. P. A. Audia Ariantari [2036268]

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Business Analytics and Operations Research

Tilburg School of Economics and Management
Tilburg University

Supervisor:
Dr. Cristian Dobre

July 8, 2021

Abstract

In this thesis, the analysis on the number of infections with the SARS-CoV-2 virus is discussed. The data contains four status of patients' which provided in the online dashboard *rijksoverheid.nl*. The chosen period is July 2020 until January 2021, as July 2020 marks the month when a large scale of testing is done and January 2021 as the end of the period since there is the start of vaccination which influences the transition from being infected to being in need of hospitalization.

In order to obtain a reliable result, the mathematical model from these data is estimated by two methods. The methods are Segmented Least Squares (SLS) and Regression in the presence of a qualitative factor. The result is carried out in Markov chain evolution to observe the situation of SARS-CoV-2 in the Netherlands over time.

It turns out that dynamic programming on multi-way choices validates that there are window shifts between states. The proportions to shift from one state to another are obtained by a centered model in general linear form. These respective proportions then translated to the transition probabilities in Markov Chain. Subsequently, the rollout of Markov Chain evolution showed that the actual condition on early pandemic days is much worse than it is shown in the online dashboard. It would have taken only four more weeks for the virus to spread to the whole Netherlands population if there was no governmental interference.

Contents

1 Introduction **5**

2 Literature Review **7**

3 Segmented Least Squares (SLS) **8**
3.1 The Notion of SLS 8
3.2 Formalizing the Notion 9
3.3 SLS: The Algorithm 9

4 Regression in Presence of a Qualitative Factor **12**
4.1 Overview of Analysis of Covariance (ANCOVA) 12
4.2 The Regression with a Qualitative Factor 12
4.3 The Model in General Linear form (GLM) 13

5 Markov Chains **15**
5.1 Introduction to Markov Chains 15
5.2 Evolution Over Time 15
5.3 Condition for Limiting and Stationary Distribution 16
5.4 The role of Markov Chains 18

6 Implementation for SARS-CoV-2 Data **19**
6.1 Introduction to the Netherlands SARS-CoV-2 Data Problem . . 19
6.2 Implementation of Segmented Least Squares 21
6.2.1 SLS Implementation on Detected Infected (DI) 21
6.2.2 SLS Implementation on Hospitalized (Hos) and Intensive Care Unit (ICU) 22
6.2.3 Window Shift between Patient Status 23
6.2.4 Attempt of SLS Implementation on Dead (DD) 26
6.3 Implementation of Regression in a Presence of Qualitative Factor 29
6.3.1 The Trends 29
6.3.2 GLM form: Estimated Lines for Each Segment 33
6.3.3 Retrieved Proportion 34
6.4 Implementation of Markov Chains 36
6.4.1 Transition Probability Matrix with Five States 36
6.4.2 Markov Chain Simulations 38
6.5 An Analysis of SARS-CoV-2 in the Netherlands on Early Pandemic Days 40
6.6 Reconstruction of March 4th 2020 until May 12th 2020 43

7 Summary, Conclusions, and Recommendations **44**

References **46**

Appendices **47**

A Appendix: Day shifts Tables **47**
A.1 DI-Hos shift on Day 7th 47
A.2 DI-Hos shift on Day 8th 48
A.3 DI-Hos shift on Day 9th 50
A.4 DI-Hos shift on Day 10th 51
A.5 DI-Hos shift on Day 11th 52
A.6 DI-Hos shift on Day 12th 53
A.7 DI-Hos shift on Day 13th 55
A.8 DI-Hos shift on Day 14th 56
A.9 DI-Hos shift on Day 15th 57
A.10 DI-Hos shift on Day 16th 58
A.11 DI-Hos shift on Day 17th 60
A.12 DI-Hos shift on Day 18th 61

Acknowledgements

I am incredibly grateful to my supervisor, Dr. Cristian Dobre, for his guidance, invaluable advice, continuous support, and patience during the writing process of this thesis.

I am profoundly grateful to my colleague, Marc van Iersel, for working with me side by side for this master's degree journey.

Finally, I would like to express my abundance of gratitude to my parents. Without their tremendous understanding and encouragement, it would be impossible for me to complete my study.

1 Introduction

Coronavirus disease, also known as COVID-19, is an infectious disease caused by the coronavirus SARS-CoV-2. The COVID-19 has become a pandemic in 2020. SARS-CoV-2 infection remains a significant challenge for the world ever since. According to the World Health Organization (WHO), people infected with the virus will experience respiratory problems scaling to mild to moderate. In most cases, infected people will not require special treatment to recover. However, older people and people with underlying medical problems are more prone to develop a major illness. Garg et al. (2020) states the most common underlying conditions were hypertension, obesity, chronic lung disease, diabetes mellitus, and cardiovascular disease. Additionally, Centers for Disease Control and Prevention CDC (2020) states pregnant and recently pregnant people are also at increased risk for severe illness.

WHO (2020) states the best way to prevent or slowing down virus transmission is by being updated with any information about the virus, the diseases it causes, and the way it spreads. As the virus primarily spreads through droplets of saliva or nose discharge of infected person, the virus could spread easily if people do not practice respiratory etiquette, e.g., coughing into a flexed elbow.

The rising number of infected people resulted in rising demand for special treatment from a hospital, whether medical assistance given in a hospital bed or intensive assistance given in an intensive care unit. If the virus transmission continues without any interference, there will be a high demand for medical assistance that could cause disruptions.

In the Netherlands, the first disease has been detected in early March 2020. The government needed to make a quick interference in order to slow down the virus transmission by creating national measures. The question now will be how fast the virus will spread if the government does not interfere and how many hospital beds and intensive care beds are needed to accommodate the infected people with a major illness.

In this thesis, one performs an analysis of the SARS-CoV-2 condition in the Netherlands. This research aims to know the number of patients infected with the SARS-CoV-2 virus and their status over time. The difficulties here lie in the assumption that has not been statistically proven yet. Therefore, this thesis aims to obtain the statistically proven assumption.

An estimation method called Segmented Least Squares (SLS) is used as the first step of analyzing SARS-CoV-2 data. SLS allows one to understand the segmentation of patients according to the trend with their respective days. There are a polynomial number of possibilities to obtained a reliable and optimal segmentation, known as Multi-way choices. Thus, dynamic programming is then used to solve this problem. Implementing SLS ables one to validate assumptions about how many days approximately patients will change their status.

Another estimation method discussed in this thesis is the study of regression in the presence of a qualitative factor. The regression model has the same mathematical model as ANCOVA. The model estimation leads to some lines, which then being tested with F-statistic to know whether these lines obtained are parallel or not. Subsequently, the proportion of patients that change their health status can be retrieved here.

These proportions are used for the next step in Markov chain simulation.

Markov chain has its unique trait that describes a step-by-step movement, which in this case is the change of patient status. Therefore, Markov chain plays an essential role in describing the process of SARS-CoV-2 patients' status at a particular time.

The goal of this thesis shows the SARS-CoV-2 condition over time in the Netherlands with a statistically proven assumption. In order to see the condition over time, Markov chain evolution and Markov simulations are performed.

The thesis outline looks as follows: First, a literature review is written in the second chapter. Second, a section about Segmented Least Squares (SLS) is given. This is followed by a section about regression in the presence of a qualitative factor. Then, a section about the implementations of the method for SARS-CoV-2 data is explained; it includes data introduction, SARS-CoV-2 analysis in early pandemic days, and the reconstruction of early days data. The latter is a section containing a summary, conclusions, and recommendations.

2 Literature Review

The data from *rijksoverheid.nl* has intrigued one to conduct an analysis of SARS-CoV-2 in the Netherlands. This online dashboard provided data on detected infected people, hospital admissions, intensive care unit admissions, deaths, and reproduction factor.

Shown in Figure 1 is the reproduction factor over time. The data showed February 17th, 2020, as the start of the period. However, there was no extensive scale testing before July 2020. There is not much information on how the reproduction factor is being estimated. The reproduction factor is crucial to know how fast the virus will spread. By understanding the reproduction factor, one could understand in which direction the condition of the SARS-CoV-2 will be.

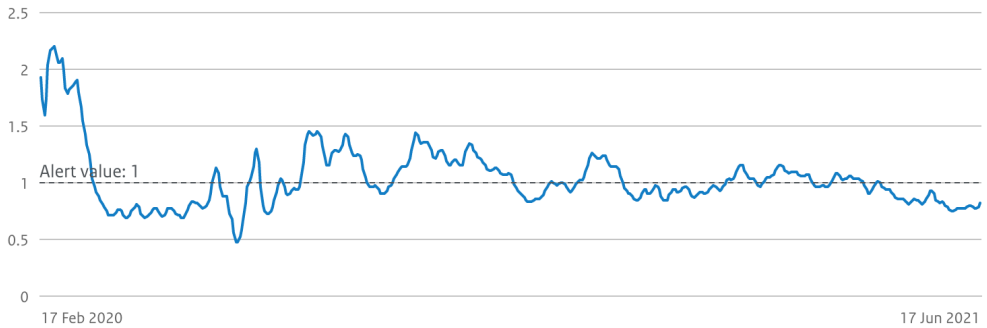

Figure 1: Reproduction Factor Over Time (Source: RIVM)

Due to the lack of information on the change of patients' status (whether being hospitalized, being in ICU, or dead), COVID-19 op de Nederlandse Intensive Cares by NICE (2021) is considered. This is the only paper in the Netherlands with specific information about the Intensive Care Unit (ICU). The data is collected throughout all ICU units in the Netherlands. NICE (2021) has provided information on the number of patients that are being treated in ICU, the number of patients that are being discharged from ICU and move to a usual hospital bed, and the number of patients that died in ICU starting from March 2020.

In this thesis, the data from *rijksoverheid.nl* is analyzed by estimation methods. The estimation gives the idea of the change of patients' status from detected infected to hospitalized and hospitalized to the intensive care unit. While the additional data by NICE (2021) specifically complements this thesis with ICU data which gives information on the proportion of patients' status change from ICU to hospitalized and from ICU to dead.

3 Segmented Least Squares (SLS)

This chapter discusses about one of the two methods that are used in this research. The Segmented Least Squares (SLS) is used to estimate the piecewise linear model parameters. As this estimation involves many possibilities to find the optimal solution, this chapter also discusses how dynamic programming plays a vital role in solving this problem.

3.1 The Notion of SLS

Suppose that the given data denoted as $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ are n points in a plane that are also the elements of a set P and suppose that $x_1 < x_2 < ... < x_n$. The goal is to find a line of best fit L (a line with minimum error) with respect to P. This error of line L is the sum of its squared distance to the points in P. The line L is defined by the equation $y = ax + b$, hence the error of L is given by:

$$Error(L, P) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \tag{1}$$

The $\hat{a}$ and $\hat{b}$ values where the line leads to a minimum error are given by:

$$\hat{a} = \frac{n \sum_i x_i y_i - (\sum_i x_i)(\sum_i y_i)}{n \sum_i x_i^2 - (\sum_i x_i)^2} \tag{2}$$

$$\hat{b} = \frac{\sum_i y_i - a \sum_i x_i}{n} \tag{3}$$

In some cases, the data points visualization might look like in Figure 2. In this case, the formulas given were not designed to cover such phenomena. In essence, any single line through the set of data points in Figure 2 could lead to terrible error. Thus, by judging the way the data points are plotted, there are approximately two lines in Figure 2 that could achieve minimal error. The same things applied to Figure 3. There are approximately three lines. The following purpose is to formalize this notion.

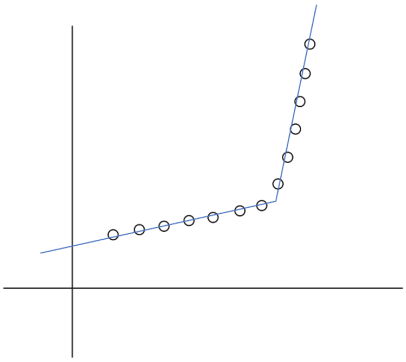

Figure 2: Two lines covers data points

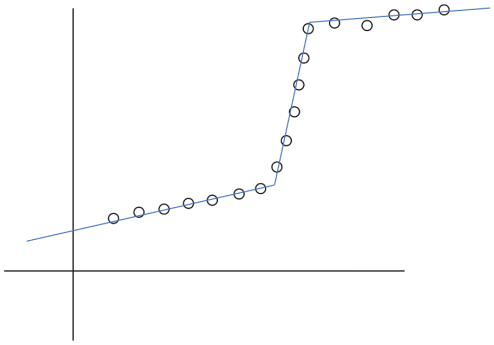

Figure 3: Three lines covers data points

3.2 Formalizing the Notion

The notion mentioned comes from the intuition of only looking at the figure. On the other hand, instead of trying to fit a single line, an arbitrary set of lines can be chosen to minimize the error. Unfortunately, this is not a suitable problem formulation. If one can fit the points to an arbitrary set of lines, then one could also fit the points perfectly by just choosing a new line passing on each pair of consecutive points. Hence, one needs a better problem formulation to fit the points well using as few lines as possible.

The better problem formulation is known as change detection. Given a sequence of data points, one wants to identify a change from one linear estimation to another (from a few points in the sequence to another). Thus, given a set of points in $P = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$ and $x_1 < x_2 < ... < x_n$. The point (x_i, y_i) is denoted as p_i. Partition set P into several subsets. Each subset, say S is of the form $\{p_i, p_{i+1}, ..., p_{j-1}, p_j\}$ for some indices $i < j$. Finally, for each S in partition of set P, compute the line that minimize error with respect to points in S. The penalty of a partition is called trade-off function defined as: $E+CL$, for some constant $C > 0$. E is the error value of the optimal line in each segment and L is the number of lines. Consequently, if one wants to minimize error (by tuning C into some low number) as much as possible then the number of segments/lines are increasing and vice versa.

The Segmented Least Squares method aims to find a partition with a minimum penalty. However, there are a polynomial number of possibilities in order to find this partition. The following subsection starts with explaining dynamic programming's ability to find a partition of the minimum penalty with a polynomial number of possibilities.

3.3 SLS: The Algorithm

Nowadays, Dynamic programming is a widely-known method to solve many optimization problems. In general, the intuition of dynamic programming is to explore all the possible solutions by working backward from the end of a problem towards the beginning, breaking large problems into smaller subproblems, then building up correct solutions to larger and larger subproblems.

As mentioned before, Segmented Least Squares method involves a polynomial number of possibilities to find the optimal solution at each step, also called

Multi-way choices. Dynamic programming plays its natural role in solving this problem.

Originally, the first segmented straight lines observation by Bellman and Roth (1969) has been elaborated by McZgee and Carleton (1970). It was continued by Kim et al. (2008) that observe on segmented line with a simulation study. Recently, Mankowski and Moshkov (2021) uses dynamic programming for multi-objective optimization.

The following observation by Kleinberg and Tardos (2005) is to find an optimal solution from Segmented Least Squares. A possible optimal partition as seen in Figure 4 can be obtained with p_i as the beginning of the single line segment and ends at the last point p_n. Accordingly, if the identity of the last segment $p_i, p_{i+1}, ..., p_n$ is known, then one could remove these points from consideration as it is already the last segment of the possible optimal partition. Hence, solving the subproblem on the remaining points $p_1, ..., p_{i-1}$ recursively will lead to the next possible segment of the possible optimal partition.

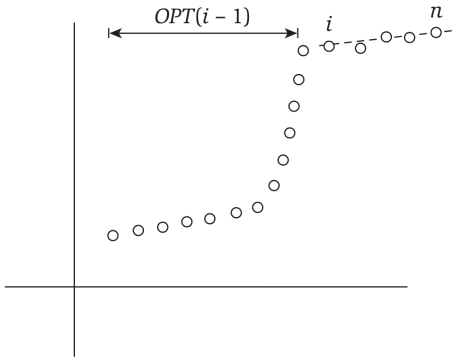

Figure 4: A possible optimal solution
(Kleinberg, 2005)

Suppose the optimum solution for the points $p_1, ..., p_i$ denoted as $OPT(i)$ and the minimum error of any line with respect to points $p_i, p_{i+1}..., p_j$ denoted as $e_{i,j}$. The observation above says the following (Kleinberg, 2005, p. 265):

1. If the last segment of the optimal partition is $p_i, ..., p_n$ then the value of the optimal solution is $OPT(n) = e_{i,n} + C + OPT(i-1)$
2. For the subproblem on the points $p_i, ..., p_j$,

$$OPT(j) = \min_{1 \le i \le j} (e_{i,j} + C + OPT(i-1)) \tag{4}$$

and the segment $p_i, ..., p_j$ is used in an optimum solution for the subproblem if and only if the minimum is obtained using index i.

Below, the algorithm is given in pseudo code.

Algorithm 1: Segmented Least Squares algorithm

Data: $n, p_1, ..., p_n, C$
Result: Near optimal solutions $OPT(i)$ in order of increasing i
Segmented Least Squares(n)
Array $M[0...n]$
Set $M[0] = 0$
for *all pairs* (i, j) *with* $i \leq j$ **do**
  Compute the least square error $e_{i,j}$ for the segment $p_i, ..., p_j$
end
for $j = 1, 2, ..., n$ **do**
  $M[j] = \min_{1 \leq i \leq j}(e_{i,j} + C + M(i-1))$
end
return $M[n]$

4 Regression in Presence of a Qualitative Factor

The second method to estimate the model parameters is the study of regression in the presence of a qualitative factor. This chapter starts with an overview of ANCOVA, which has the same mathematical model as the regression in this research. The second part discusses this mathematical model as the regression in the presence of a qualitative factor. Finally, the last part discussed the model in General Linear (GLM) form.

4.1 Overview of Analysis of Covariance (ANCOVA)

ANCOVA is a technique invented by R.A. Fisher. The first description of ANCOVA is in his book "Statistical Methods for Research Workers" in 1930, which later improved in 1935. This technique is first applied to increase the precision of comparison of treatments by H.G. Sanders, which Fisher advised in 1930.

ANCOVA itself is a part of the General Linear Model (GLM). Typically, GLM refers to conventional linear regression models for a continuous response variable given continuous and/or categorical predictors. Consequently, one has an ANOVA when the model has no continuous factors. If the model has no categorical factors, then it is a regression. At last, one has General Linear Model (GLM) when the model has both continuous and categorical factors. ANCOVA can be used to include both of these factors.

Let x denote the covariate (continuous predictor) which is included as explanatory variable. Commonly, a covariate is centered: $z = x - \bar{x}$. Let β_0 denote the intercept which is the expected yield of the reference treatment and τ_t being the parameter corresponding to the qualitative factors. Let τ_t denote the fixed effects which is the difference in expected yield between treatments. Subsequently, let β_1 denote the random effects induced by the quantitative factor and let ϵ_{tb} denote the random error term. Let $\mathcal{T}$ denote the number of treatments and $\mathcal{B}$ denote the number of subjects for each treatment. The ANCOVA mathematical model is given by:

$$\begin{aligned} y_{tb} &= \beta_0 + \tau_t + \beta_1 z_{tb} + \epsilon_{tb} && t = 1, ..., \mathcal{T} \\ &= \beta_0 + \tau_t + \beta_1 (x_{tb} - \bar{x}) + \epsilon_{tb} && b = 1, ..., \mathcal{B} \end{aligned} \tag{5}$$

There are two steps that initially lead R.A. Fisher to introduce ANCOVA Model. First, he performed regression of y on z. He then changed the response y into $(y - \hat{\beta}_1 z)$ for correcting z. Second, he performed ANOVA on these corrrected observations: $y_{tb} - \hat{\beta}_1 z_{tb} = \beta_0 + \tau_t + \epsilon_{tb}$. At last, the model in (5) is obtained by moving $\hat{\beta}_1 z_{tb}$ from left hand side to the right hand side.

There are several assumptions for ANCOVA model which are independence between error terms, normality of error terms, equal variance of error terms, linear relationship between response y and covariate x, and covariate x does not depend on the treatments.

4.2 The Regression with a Qualitative Factor

As mentioned by Kutner et al. (2005) about generalization of covariance model, there are two points of view towards ANCOVA model. The first one is there is interest in the treatments. This is what people usually have in mind for

analysis of covariance. The covariate x is introduced to increase precision of comparison between treatments. The second one, that is the purpose of this research, is a specific interest in the relationship between y the response variable and x explanatory variable. This is called the regression in the presence of a qualitative factor. In this part, variable x is of interest as well and not merely there to increase the precision of comparison between treatments.

Suppose that the relationship between response y and a covariate x (or several covariates $x_1, x_2, ...$) is linear and suppose that parameters of covariates may differ between treatments. Then, product terms of covariates and dummy variables are introduced here. Essentially, these product terms represent interactions between factors and covariates. Subsequently, the model assumptions are almost same with ANCOVA model assumptions given before. The only difference is the last assumption, covariate x does not depend on the treatments, is no longer applied as one is not focusing on analysis of covariance.

4.3 The Model in General Linear form (GLM)

Fitting the model in equation (5) is done to match the data. Notice that there is only β_1 there, which means one is imposing an equal slope, i.e., parallel lines. In general, the model in (5) can be designed to allow different slopes. The GLM approach is made to allow different slopes shown in Rutherford (2001) is presented here. The following model is in General Linear (GLM) form, which given by:

$$y_{ij} = \beta_0 + \tau_i + \beta_1 x_{ij} + \lambda_i x_{ij} + \epsilon_{ij} \qquad \begin{matrix} i = 1, ..., I \\ j = 1, ..., J \end{matrix} \tag{6}$$

The model in GLM form given in equation (6) allows different slopes for different treatment as interaction terms are present. Let I denote the number of treatments and let J denote the number of subject for each treatment. Thus, this model in GLM form allow the difference in intercepts as well as the difference in slopes.

The difference in intercepts is the difference in expected yield between treatments provided by τ_i. If the value of τ_i is 0 then that particular treatment is a reference treatment. While the other treatment i will have some value for τ_i accordingly. Subsequently, the difference in slopes which is the difference of the random effects induced by the quantitative factor for each treatment is given by λ_i. This means the slope of all treatments will no longer be the same, in other words it is no longer imposing parallel lines. The value of λ_i is 0 when the respected treatment is reference treatment while the other treatment will have this value accordingly.

Regression in presence of a qualitative factor model is adopted to know whether the relationship between response y and a covariate x (or several covariates $x_1, x_2, ...$) is affected by the treatments. The null hypothesis corresponding to this regression model present that there is no different slope between treatments (7), while the alternative hypothesis present that at least one treatment's slope is different (8).

$$H_0 : \lambda_1 = \lambda_2 = ... = \lambda_{i-1} = 0 \tag{7}$$

$$H_\alpha : \lambda_i \neq 0 \qquad \exists i = 1, ..., I-1 \tag{8}$$

The F-statistic compares the mean square of the fixed factor with the mean square of the error term. The larger the F-statistic indicates a high probability of a difference between the slopes of the treatments. The F-statistic follows an F-distribution with $(I-1)$ and $(I-1)$ x $(J-1)$ degrees of freedom. The α denoted the chosen significance level. The typical value of α is 5%, which establishes a 95% confidence level. This indicates that there is a probability of 5% to find an outcome of the F-statistic larger than the critical value, given that the null hypothesis is true.

Nowadays, modern statistical software condenses F-test by providing the p-value. The p-value here is the probability of getting an F-statistic even greater than what one observes. In other words, the more F-statistic values, the lesser becomes the p-value. Hence, the decision rule is if the p-value obtained is less than α, then Reject H_0 and Accept H_α.

Therefore, first, imposing equal slope (parallel lines), hence parameters λ_i are set to be zero. Then, the F-test is performed. This step is to investigate whether it is suitable to impose equal slope or not. When the model is applied and the F-test rejects the null hypothesis, it is known that there is a difference between the slopes of the treatments.

5 Markov Chains

This chapter start with an introduction to Markov chains and its evolution over time. The model formulation of Markov chain over time will be used to tackle the problem of this research. Therefore, only Markov theory related to Markov chain over time roll-out that is relevant for the problem, will be introduced.

5.1 Introduction to Markov Chains

Let X_n denote values in each time period n of a process and let $\{X_n, n = 0, 1, 2, ..., \}$ be a stochastic process that takes on a finite values of non-negative integers. A Markov chain is a special stochastic process that describes movement step-by-step through a number of states. Each state is referring to a situation in the process at a certain point in time n. Hence, a state describes the current position of the process. All the situations in the process are in the state space $\mathcal{S}$. The process is always in one state of the state space $\mathcal{S}$ at any point in time. Furthermore, the distribution of the next state, given the past and current states, only depends on the current state and not on the past states. This is also known as the *Markov property*. Thus, one needs to determine the initial state in order to describe a Markov chain. Since Markov chain is a stochastic process, there is a probability for certain transition between one state to another in the process. Suppose that whenever the process is in state i, there is a probability p_{ij} that it will be in state j. Then, the Markov property for a Markov chain is given by:

$$\begin{aligned} &Pr\{X_n = j_n | X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, ..., X_1 = i_1, X_0 = i_0\} \\ &\quad = Pr\{X_n = j_n | X_{n-1} = i_{n-1}\}, \quad i, j \in \mathcal{S}, \quad i, j \geq 1, \quad n \geq 0 \end{aligned} \tag{9}$$

Let matrix P be the matrix with p_{ij} as its elements which is known as the *transition probability matrix*. Matrix P is a stochastic matrix by definition, as a consequence, all elements of P are non-negative $p_{ij} \geq 0$ with $i, j \geq 0$ and its rows sums up to one, i.e. $\sum_{j=0}^{\infty} p_{ij} = 1$ with $i = 0, 1, 2, ...$ The value of p_{ij} correspond to a one-step transition probability that the process will make a transition to state j, given state i as a current position, is given by:

$$p_{ij} = Pr\{X_n = j | X_{n-1} = i\} \tag{10}$$

5.2 Evolution Over Time

Previously, p_{ij} were already defined as the one-step transition probabilities. Now, let p_{ij}^n be the n-step transition probabilities. One should understand how the chain evolves over time in order to study the performance characteristics of Markov chains. Accordingly, one should look at the n-step transition probabilities p_{ij}^n to comprehend the process over time.

The Chapman-Kolmogorov equations establish a method to compute these n-step transition probabilities out of the one-step transition probabilities, by means of the first law of total probability and the Markov property. The formal

derivation as described by Ross (2014) is given in expression (11) as follow:

$$\begin{aligned} p_{ij}^{n+m} &= Pr\{X_{n+m} = j | X_0 = i\} \\ &= \sum_{k=0} Pr\{X_{n+m} = j, X_n = k | X_0 = i\} \\ &= \sum_{k=0} Pr\{X_{n+m} = jX_n = k, X_0 = i\} Pr\{X_n = k | X_0 = i\} \\ &= \sum_{k=0} Pr\{X_{n+m} = jX_n = k\} Pr\{X_n = k | X_0 = i\} \\ &= \sum_{k=0} p_{kj}^m p_{ik}^n, \qquad i, j \in \mathcal{S}, \quad \forall n, m \geq 0 \end{aligned} \tag{11}$$

Expression (11) is obtained by determining the probability that state i as initial state will eventually reach state j in $(n+m)$-transitions by conditioning on the position reached after m steps. Consequently, the process will no longer depend on past behavior in the evolution of the process over time once the process has reached state k with m-transitions. The future behavior then only depends on the current state k. Thus, the probability that the process will be in state j after $(n+m)$-transitions is equal to the summation of the probabilities of all intermediate state k.

Let $P^{(n)}$ denote the matrix of n-step transition probabilities p_{ij}^n then expression (11) also state that:

$$P^{(n+m)} = P^{(n)} P^{(m)} \tag{12}$$

$$P^{(n)} = P^{(n-1+1)} = P^{(n-1)} P = P^n \tag{13}$$

From expression (12), in particular, $P^{(2)} = P^{(1+1)} = PP = P^2$. Then, expression (13) is obtained by induction. Therefore, the n-step transition probability matrix can be obtained by multiplying the one-step transition probability matrix P by itself n times.

5.3 Condition for Limiting and Stationary Distribution

State j is said to be *accessible* from state i if $p_{ij}^n > 0$ for some integer $n > 0$. This implies that state j is accessible from state i if and only if starting from the initial state i, the process will ever enter state j. The two states i and j are said to *communicate* when these states are accessible to each other. This indicate that if state i communicates with state j, then state j communicates with state i. Also, if state i communicates with state j and state j communicates with state k, then state i communicates with state k. Thus, state i is accessible from state k and vise versa. If all states in space state $\mathcal{S}$ communicate with each other then this chain is said to be *irreducible*. The state that only communicates with itself and the probability of staying in that state equals to 1 is called *absorbing* state.

For any state i, let f_i denote the probability that the process will reenter state i given state i as a starting state. If $f_i < 1$, then state i is a *transient* state. Moreover, if $f_i = 1$, then state i is a *recurrent* state. As the process always needs to be in some state, every Markov chain needs to have at least one

recurrent state since the fact that if all the states are transient, then the process will be in no state. When the expected number of transitions that state i takes to return to itself is finite then state i is said to be *positive recurrent* and it said to be *null recurrent* when it takes infinite transitions.

$$f_i = \sum_{\forall k \in R^+} Pr\{X_{n+k} = i | X_n = i\}, \qquad i \in \mathcal{S}, \quad n \geq 0 \tag{14}$$

State i of a Markov chain is said to have period d if that state i will be revisited whenever n, number of transitions, is not divisible by d. That implies with $d > 1$ then a state is *periodic*. However, when $d = 1$ then a state is *aperiodic*.

A Markov chain is said to have a limiting distribution if the chain staisfies 3 conditions which are *irreducible*, *aperiodic*, *positive recurrent*. Also, Pinsky and Karlin (2011) described that a limiting distribution, when it exists, is always a stationary distribution, but the converse is not true. There may exist a stationary distribution but no limiting distribution.

When a Markov Chain is *irreducible*, *aperiodic*, *positive recurrent* then there is a probability $\pi_i^{(n)}$ that the process is in state i at time n. This probability will converge to a limit π_i as n goes to infinity, regardless the initial state. Thus, letting π_j denote the long-run proportion of time that the chain is in state j. Formally stated in theorem described by Ross (2014) as follow:

Theorem Let $\{X_n, n = 0, 1, ...\}$ be an irreducible, aperiodic, positive recurrent Markov chain. Then exists $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j \ \forall j \in \mathcal{S}$. This limit is independent of the value of $i \in \mathcal{S}$. Then $\{\pi_j, j \in \mathcal{S}\}$ is the unique non-negative solution of the equation (15) and equation (16).

$$\pi_j = \sum_{i \in \mathcal{S}} \pi_i p_{ij}, \qquad j \in \mathcal{S}, \tag{15}$$

$$\sum_{j \in \mathcal{S}} \pi_j = 1 \tag{16}$$

Equation (15) can also be written as equation (17) where π is the row of limiting probabilities π_j and P the matrix of one-step transition probabilities.

$$\pi = \pi P \tag{17}$$

The set $\{v_j, j \in \mathcal{S}\}$ is called the stationary distribution of a Markov chain. A stationary distribution is such a distribution v that if the distribution over states at step k is v, then the distribution over states at step $k + 1$ is v. Therefore, a stationary distribution can also be expressed by:

$$v = vP \tag{18}$$

Notice that this expression on equation (18) is similar to equation (17). The only difference is v_j is not uniquely defined.

Consequently, the statement from Pinsky and Karlin (2011) can be restated as follow:

1. The limiting distribution of a regular Markov chain is a stationary distribution.

2. If the limiting distribution of a Markov chain is a stationary distribution, then the stationary distribution is unique.

In practice, limiting and stationary distribution allows one to evaluate the probability of ending up in a particular recurrent class, the (mean) time until entering one of the recurrent classes, and the long-run proportion of time spent in each states. This particular research is focusing on evaluating the long-run proportion of time spent in each of its states.

5.4 The role of Markov Chains

Markov Chains plays an essential role in this research. As its unique trait describes a step-by-step movement, Markov chains can be used in many studies to understand a process flow. For instance, in the SARS-CoV-2 process, the Markov chain becomes incredibly valuable in describing the process of SARS-CoV-2 patients' status at a particular time.

The transition probability matrix is the one that stores valuable information on how the process moves from one state to another. This matrix also gives information on how each state would behave. Additionally, The n-step transitions probabilities have allowed one to comprehend how the chain evolves. Thus, one could also see how the process will be in the long run.

In this research, one could observe how patient status change over time and how likely this patient will be in a critical condition or even passing away. The implementation of Markov chains for an analysis of SARS-CoV-2 will be explained in section 6.4.

6 Implementation for SARS-CoV-2 Data

In this chapter, the Netherlands SARS-CoV-2 data is introduced first. Second, the implementation of the methods is explained. The two models in this chapter are both regressions. In general, one has sets of the Netherlands SARS-CoV-2 data and is determined to search for possible patterns to fit the data with the slightest error. These implementations aim to estimate the model parameter by employing Segmented Least Squares, which leads to a piecewise linear function or lines in the Regression in Presence of a Qualitative Factor with levels. Third, the implementation of Markov Chain is explained. Fourth, the analysis of SARS-CoV-2 in the Netherlands on early pandemic days is presented. In the latter section, the reconstruction of early days pandemic if there were no governmental restrictions is explained.

6.1 Introduction to the Netherlands SARS-CoV-2 Data Problem

The Netherlands SARS-CoV-2 data are obtained from *rijksoverheid.nl*. There are four types of the Netherlands SARS-CoV-2 data used which are Detected Infected (DI) people data, Hospitalized (Hos) people data, Intensive Care Unit (ICU) data, and Dead (DD) people data. These data then become the states of patients for Segmented Least Squares method and Markov Chain, while it becomes qualitative factors for Regression in Presence of a Qualitative Factor.

Initially, these four states/qualitative factors are considered for the parameter estimations of this research's model by utilizing both the segmented least squares method and the least square method (The study of Regression in the Presence of a Qualitative Factor). However, the Dead (DD) data is no longer being taken into account in the implementation of estimation methods due to the high volatility of data that jeopardizes the estimation performed. This rapid and unpredictable change happens because of many factors that are outside the scope of this research. One of the many factors is that patients who passed away (state DD) do not necessarily come from ICU beds. In some cases, the patients that passed away could come from detected infected status or hospitalized status.

The following figures are the four types of data shown in a histogram with the line representing the seven-day average. For this research, the starting date is July 7th, 2020 until January 31st, 2021 for Detected Infected (DI). This period is chosen because starting from July 7th, 2020, there is a visible effect of the virus since a large scale of testing is implemented. In contrast, the end period is chosen at January 31st, 2021 because there is a beginning of vaccination which influences the transition from being infected to being in need of hospitalization. The starting and the end date for the other three types are adjusted according to the estimation of window shift between patient status that will be discussed in Section 6.2.

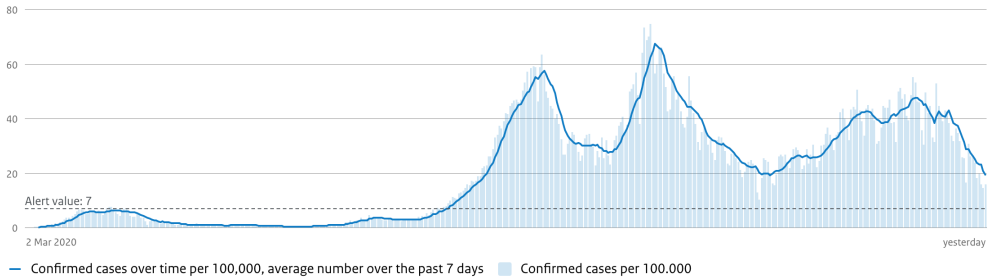

Figure 5: Number of Detected Infected over time (Source: RIVM)

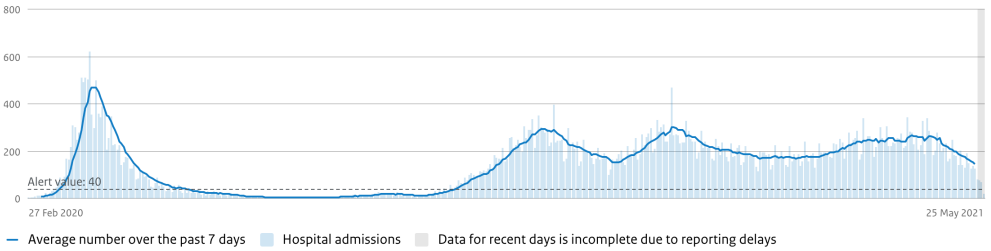

Figure 6: Hospital Bed Admissions over time (Source: RIVM)

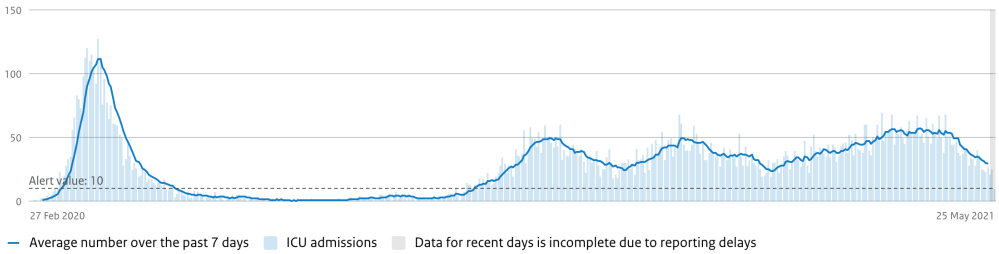

Figure 7: Intensive Care Unit Admissions over time (Source: NICE via RIVM)

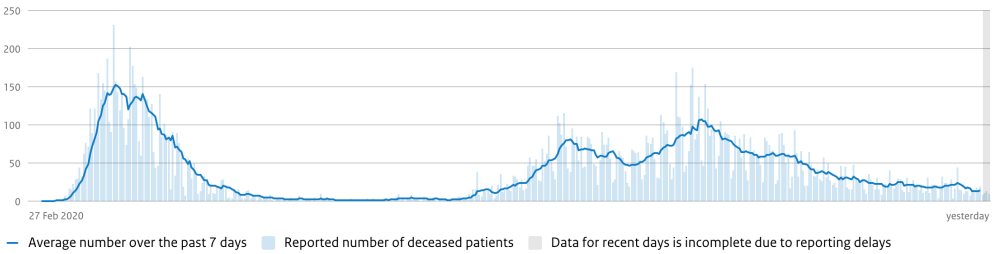

Figure 8: Deaths over time (Source: RIVM)

6.2 Implementation of Segmented Least Squares

The Segmented Least Squares method is implemented for each type of the Netherlands SARS-CoV-2 data. This implementation is performed using python 3.7. SLS method implementation will lead to piecewise lines (segments) that correspond to particular dates.

6.2.1 SLS Implementation on Detected Infected (DI)

First of all, the SLS method is implemented to Detected Infected (DI) data which runs from July 7th, 2020 until January 31st, 2021. As described in Algorithm 1 from section 3.3, one can obtain a dynamic programming solution using precomputed results which are done by getting minimum error over possible start indices for each end index, then backtrack to get segment and coefficients. After tuning the penalty factor (C), one discovers seven segments when the penalty factor is tuned to one ($C = 1$), five segments when $C = 2, 3, 4$, and four segments when $C = 5$ for DI data. The estimation shown in Figure 10, with $C = 2, 3, 4$, gives desired patterns as the estimation of the line are not overfitting as shown in Figure 9 and also not underfitting as shown in Figure 11.

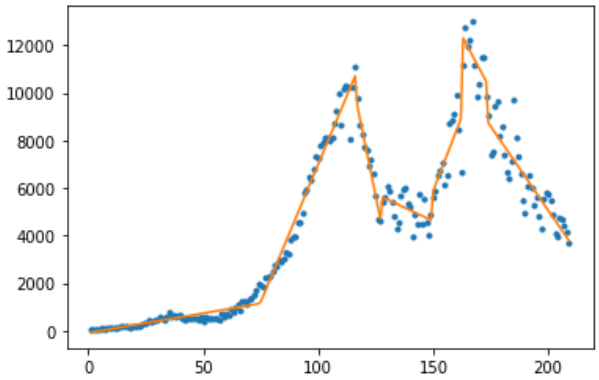

Figure 9: SLS with penalty factor C=1 for DI data

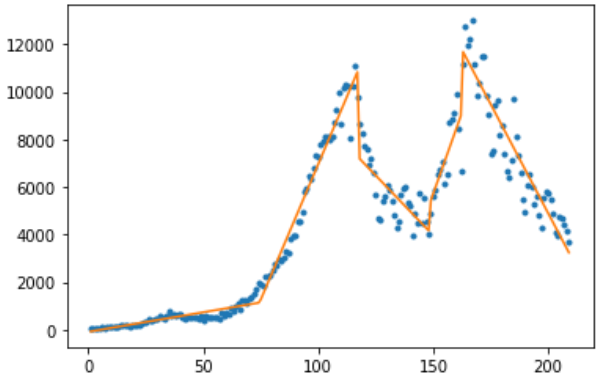

Figure 10: SLS with penalty factor C=2 for DI data

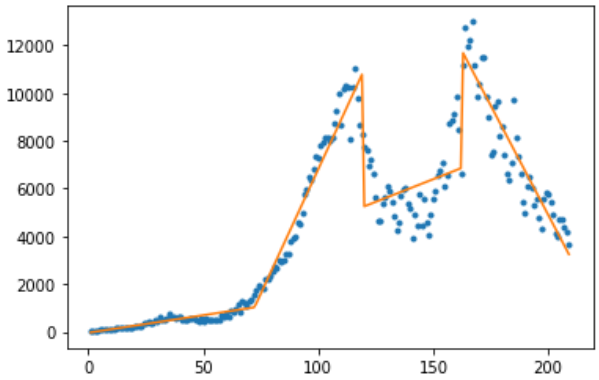

Figure 11: SLS with penalty factor C=5 for DI data

6.2.2 SLS Implementation on Hospitalized (Hos) and Intensive Care Unit (ICU)

Secondly, implementing SLS method on Hospitalized (Hos) data and Intensive Care Unit (ICU) data. There are three choices of penalty factor ($C = 2, 3, 4$) that resulted in five segments. Thus, one implement the SLS method on Hos data with penalty factor $C = 2$ first. It appears that fitting Hos data with SLS $C = 2$ resulted in a very steep kink, shown in Figure 12. This kink shows that there is a segment that only contains one observation indicating overfitting. Subsequently, one implements SLS with another penalty factor ($C = 3$) to fit Hos data. As it is shown in Figure 13, the estimation using SLS with $C = 3$ gives desirable patterns. Consequently, SLS with the same penalty factor $C = 3$ is implemented for DI, Hos, ICU consistently. Figure 14 shown the SLS implementation on ICU data.

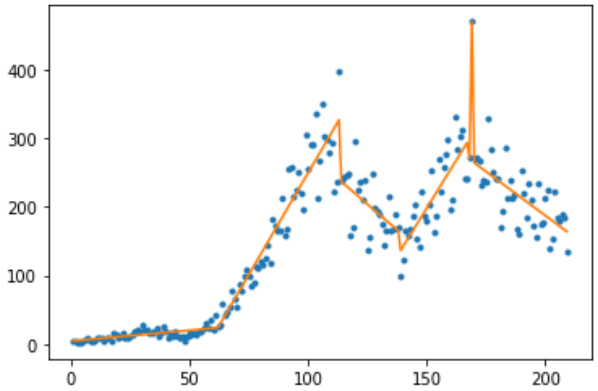

Figure 12: SLS with C=2 for Hos data

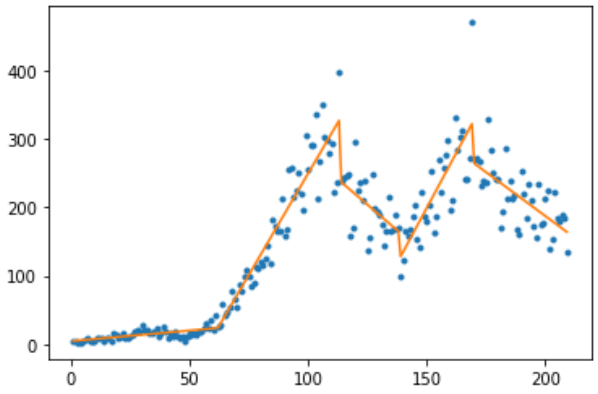

Figure 13: SLS with C=3 for Hos data

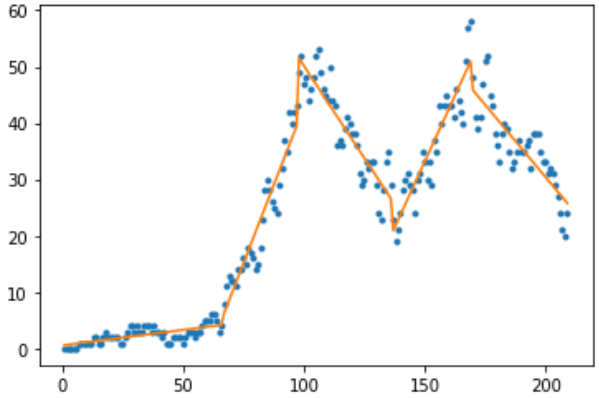

Figure 14: SLS with C=3 for ICU data

6.2.3 Window Shift between Patient Status

Detected infected people spend several days at home before being admitted to a hospital bed (from state DI to state Hos). The same situation also applies to hospitalized people who spend several days in a usual hospital bed before they are needed to be moved to an ICU bed (from state Hos to state ICU). These several days in between one state before moving to another state are called window shifts.

Detected Infected people usually have seven until eighteen days window to shift to Hospital bed while Hospitalized to ICU usually has four until nine days in between. The window shifts would vary between states. One is implementing SLS from the shift of DI to Hos and the shift of Hos to ICU data with varying windows in between. Since the window shift from DI to Hos has twelve possible days and the window shift from Hos to ICU has six possible days, there are 72 possible combinations of the implementation of SLS method.

For instance, take one possible combination of the SLS implementation. After being detected infected for seven days, the detected infected people shift to a hospital bed is written as DI-Hos: Day 7th. So, given the start date of detected infected is July 7th, the first day of state Hos here is July 13th. Subsequently, Hos-ICU: Day 4th means the hospitalized people shift to ICU on day fourth. So, given that the start date of Hos is July 13th, then the first day of state

ICU is July 16th. Consequently, the start date of combination 7-4 is July 7th, July 13th, July 16th, for DI, Hos, ICU, respectively. The end dates are adjusted respectively in order to have the same amount of observations in each state.

From this point onwards, the days in this research are called *generic days* because the start/first day of one state is not the same as the other state. Also, for the simplicity of writing the window shift combinations, starting from this point onward, DI-Hos: Day 7th and Hos-ICU: Day 4th will be written as combination 7-4. The same goes for the rest of the combinations.

Table 1: DI-Hos: Day 7th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-61	1-66
Segment 2	75-116	62-112	67-96
Segment 3	118-147	114-137	98-135
Segment 4	149-161	139-168	137-168
Segment 5	163-209	170-209	170-209

For example, as given in Table 1 for combination 7-4, the numbers shown are the generic days. The other 71 possible combinations can be seen in Appendix A. One calculates the intersection of the generic days within the same segment throughout different states. The generic days' intersection is shown as follows; Segment 1 (Day 1-61), Segment 2 (Day 75-96), Segment 3 (Day 118-135), Segment 4 (Day 149-161), and Segment 5 (Day 170-209). This calculation of generic days intersection is carried out to the other 71 combinations.

The best window shift combination is the one that has minimum mismatch from generic days intersection. For instance, in combination 7-4, the generic days' intersection is Day 1-61. Thus, there will be thirteen days mismatches for state DI, zero mismatch for state HI, and five mismatches for state ICU. In total, segment 1 of combination 7-4 has eighteen mismatches. Afterward, segment 2 has 57 mismatches, segment 3 has 38 mismatches, segment 4 has 36 mismatches, and segment 5 has seven mismatches. Consequently, combination 7-4 has 156 mismatches in total.

The result of the window shift mismatches of all possible combinations is given in Table 2. From this Table 2, there is a decreasing trend as the window shifts from DI to Hos and Hos to ICU are also decreasing. Therefore, one decides to investigate further the estimation with other combinations with lower window shifts. Therefore, three new combinations (combination 7-1, combination 7-2, and combination 7-3) are observed. The result of these mismatches of three new possible combinations is given in Table 3.

Table 2: Day shift mismatch of all possible combinations

		DI-Hos											
		7	8	9	10	11	12	13	14	15	16	17	18
Hos-ICU	4	156	162	168	174	180	186	192	198	210	222	234	246
	5	162	168	174	180	186	192	198	207	219	228	240	252
	6	168	174	180	186	192	198	207	216	228	237	249	261
	7	174	180	186	192	198	207	216	225	237	246	258	270
	8	180	186	192	198	207	216	225	234	246	255	267	279
	9	186	192	198	204	213	222	231	240	252	264	276	288

Table 3: Mismatch on 3 new possibilities

		DI-Hos
		7
Hos-ICU	1	150
	2	150
	3	153

From Table 3, one chooses either combination 7-1 or combination 7-2 since these combinations resulted in the smallest mismatches. The combination 7-2 is chosen by empirical reason. RIVM (2020) states that it will usually takes five or six days before someone develop symptoms if someone is infected with the SARS-CoV-2 virus. Then, when people have symptoms, they usually do not take the SARS-CoV-2 test right away. The reason could be that the symptoms are similar to common flu, so people tend to disregard it, or it simply takes time to wait for an appointment for them to get tested. By the time they get tested, it also takes time to get the result. It means there are many days passed before a person becomes Detected Infected.

As a result, it also affected the window shift of patient status from DI to Hos. This shift is the expected value from the estimated line obtained by employing Segmented Least Squares. According to the estimation performed here, it turns out that the shift is only seven days on average.

Additionally, patients tend to undergo some medical assistance needed in Hospital bed first; only if they become critical, they move to the ICU bed. The statement before is the empirical reason why the shift from Hos to ICU happened on average in day second. Therefore, combination 7-2 is chosen instead of combination 7-1. The following are the retrieved days from the estimation:

1. Segment 1: Day 1-61
2. Segment 2: Day 75-98
3. Segment 3: Day 118-137
4. Segment 4: Day 149-161
5. Segment 5: Day 172-209

To sum up, the best combination is when there are seven days on average for the patients to change their status from detected infected to be hospitalized.

Also, it takes approximately two days for the patients to end up in an ICU bed from a hospital bed. The retrieved start dates for combination 7-2 are July 7th, July 13th, July 14th, respectively, for state DI, state Hos, state ICU. These retrieved days are obtained by tuning $C = 3$ in SLS for combination 7-2 and are kept to be later used for the subsequent implementation of Regression in a Presence of Qualitative Factor.

6.2.4 Attempt of SLS Implementation on Dead (DD)

SLS method is now implemented on DD data. As mentioned in the last two sections, the penalty factor used for the previous three states is $C = 3$. The same penalty factor is applied to keep the consistency. According to NICE (2021), the patients tend to spend between fifteen until twenty days in ICU before passed away, i.e., move to state DD.

Therefore, the first attempt is to applying SLS with $C = 3$ on combination 7-2-15 (DI-Hos: Day 7th, Hos-ICU: Day 2nd, ICU-DD: Day 5th), which corresponds to the start date at July 28th, 2020 for DD state. However, with this penalty factor, SLS still overfit the DD data shown in Figure 15. It is also seen that the DD data has high volatility that makes the lines estimations obtained from SLS seem to overfit the data, despite tuning the same level of penalty with the other three states.

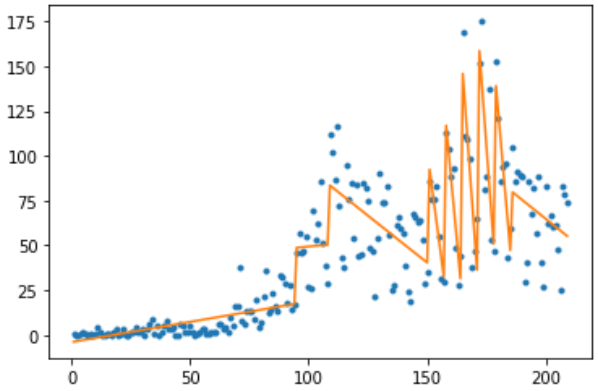

Figure 15: SLS with C=3 for DD data

Another problem with DD data is that the number of people who died is way higher than the number of people in ICU. This statement contradicts the fact that, in the Netherlands, the ICU beds are never full, which means every critical patient will undoubtedly go to ICU bed first before passing away. This problem also implies that state DD could come from any other state, either straight from DI or Hos. Hence, this implication becomes another reason why including DD data could jeopardize the estimation calculation.

Since including DD data on estimation is no longer an option, one has found the existed study that gives information about ICU beds in the Netherlands. This study is Covid-19 op de Nederlandse intensive cares by NICE (2021). The essential information regarding state ICU that is needed in this research is already provided in this study.

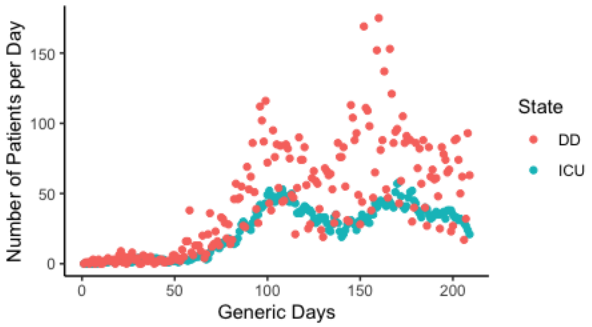

Figure 16: State ICU and State DD for the whole observed period

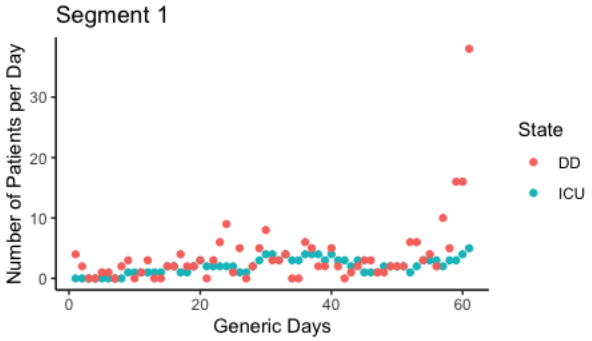

Figure 17: Segment 1 of State ICU and State DD

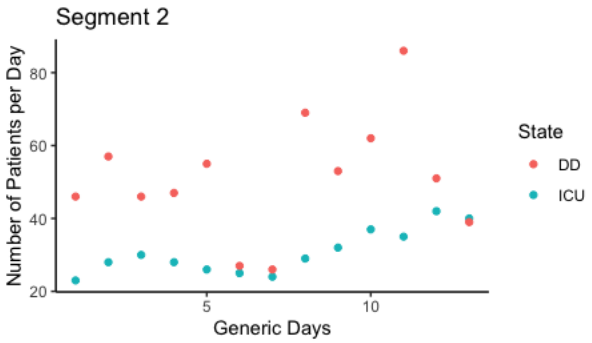

Figure 18: Segment 2 of State ICU and State DD

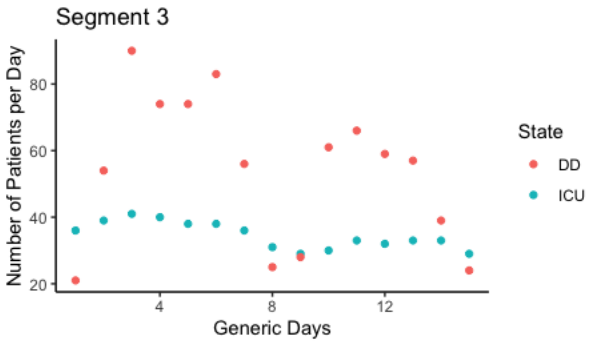

Figure 19: Segment 3 of State ICU and State DD

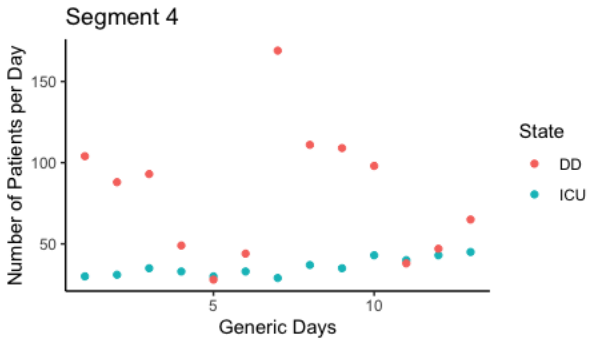

Figure 20: Segment 4 of State ICU and State DD

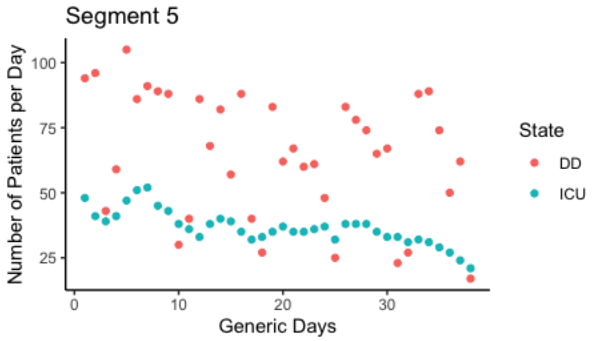

Figure 21: Segment 5 of State ICU and State DD

Figure (16) shown that DD observations are all above ICU observations. Closer look on DD observations that are always above ICU observations on each segment can be seen in Figure (17) until Figure (21).

To sum up, including DD data for this research model's parameter estimation by SLS method will lead to two problems which are:

1. The volatility of DD data forces one to tune the penalty factor to a greater number, allowing SLS to tolerate more error in a trade for fewer segments. In this case, from nine segments to five segments. Hence, besides the inconsistency of penalty tuning, this volatility compels one to tolerate more error on segment estimation in SLS, which is not the aim of this research.

2. The number of patients on DD data is greater than ICU data, which implies that patients who ended up in state DD could come from any other state. It could be from DI-DD or Hos-DD or ICU-DD. Hence, DD data is better not to be included for SLS estimation.

Since the DD data is compulsory to this research and estimating it by oneself has a risk of jeopardizing the entire estimating calculation, one decided to take a proportion from the existed paper mentioned earlier in this section. The information from NICE (2021), COVID-19 op de Nederlandse Intensive Cares, has provided this research the missing link from state ICU to state DD. This paper also provides extra information about state ICU to state Hos shift that will be explained in Section 6.4.

6.3 Implementation of Regression in a Presence of Qualitative Factor

Regression in presence of a qualitative factor model is adopted to know whether the relationship between the number of patients (y) and days (x) is affected by the states. In this regression, there are two kinds of factors which are quantitative factor (the generic days) and qualitative factors (the three states DI, Hos, ICU). The estimation then leads to three lines from each segment in Regression in Presence of a Qualitative Factor with three levels. After these three lines are obtained, one can retrieve the proportion of people that move from one state to another.

6.3.1 The Trends

Now, one wants to see whether the chosen model in (5) matches the Netherlands SARS-CoV-2 data. Let β_0 denote the intercept which is the expected yield of patients for the reference state and let τ_s being the parameter corresponding to the qualitative factors. τ_s denote the fixed effects which is the difference in expected yield between states. Next, let β_s denote the random effects induces by the generic days for each state and let ϵ_{sp} denote the random error term. Let S denote the number of states and D denote the quantitative factor which in this case is the number of days, the regression model is given by:

$$y_{sd} = \beta_0 + \tau_s + \beta_s z_{sd} + \epsilon_{sd} = \beta_0 + \tau_s + \beta_s(x_{sd} - \bar{x}) + \epsilon_{sd} \qquad \begin{matrix} s = 1, ..., S \\ d = 1, ..., D \end{matrix} \quad (19)$$

These are the 5 segments retrieved from section 6.2.3:

1. Segment 1: Day 1-61
2. Segment 2: Day 75-98
3. Segment 3: Day 118-137
4. Segment 4: Day 149-161
5. Segment 5: Day 172-209

Figure 22 until Figure 26 shows the data points for the three states DI, Hos, and ICU. One is able to see the trend of each segment in a closer look. DI data points are always on the top of the other two states (Hos and ICU) because essentially, the number of Detected Infected people is greater than Hospitalized people and people in the Intensive Care Unit. It also can be seen that Hos data points are a little bit above the ICU data points. Thus, as it is known, DI $>$ Hos $>$ ICU.

Figure 22 shows segment 1, which presents an incline from Generic Day 1-61. In comparison, Figure 23 shows an even steeper incline from Day 75-98, which appears in segment 2. Notice that the incline in segment 2 starts from approximately 2000 patients until 6500 patients. While in segment 1 the incline starts from 0 until approximately 800 patients. Figure 24 and Figure 25 show that there is a decline in segment 3, but then it starts to have another incline in segment 4. For the last segment, Figure 26 shows a trend of steep decline.

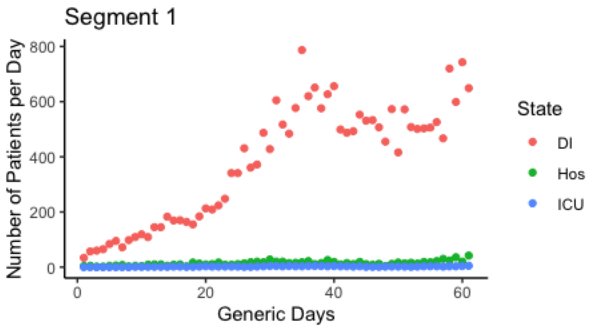

Figure 22: Segment 1 of DI, Hos, and ICU

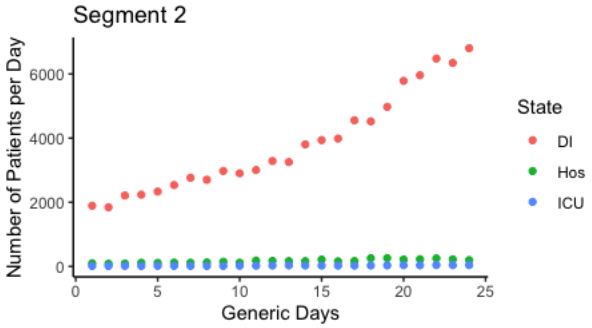

Figure 23: Segment 2 of DI, Hos, and ICU

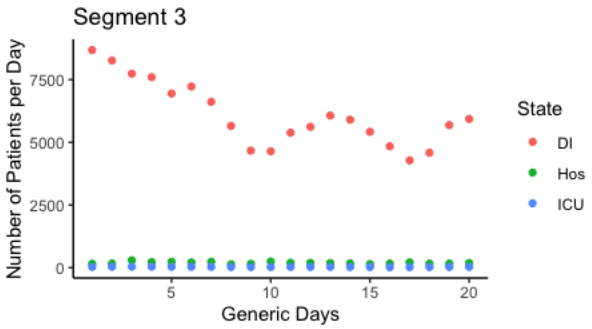

Figure 24: Segment 3 of DI, Hos, and ICU

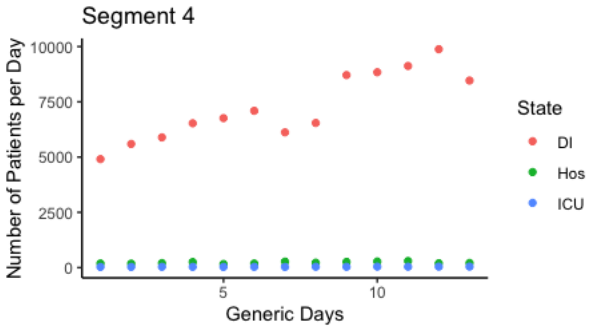

Figure 25: Segment 4 of DI, Hos, and ICU

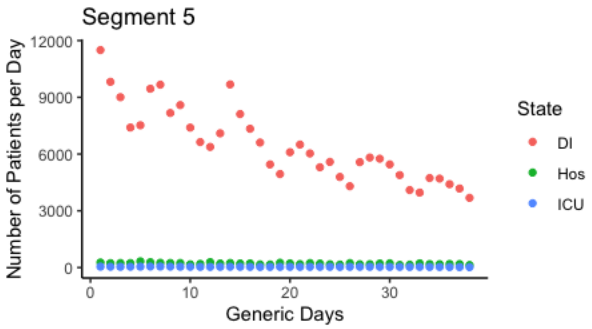

Figure 26: Segment 5 of DI, Hos, and ICU

Figure 27 until Figure 31 show the line estimations for the 3 states within one segment in one scale. This model is fitted to data through least square method. The model that is used here is a centered model, i.e., notice that $z_{sd} = x_{sd} - \bar{x}$. These figures confirmed that imposing an equal slope is not suitable for this research. Thus, the model in General Linear form, which allows different slopes, will be discussed in the next section. There will be an F-test to confirm whether imposing an equal slope is suitable statistically or not.

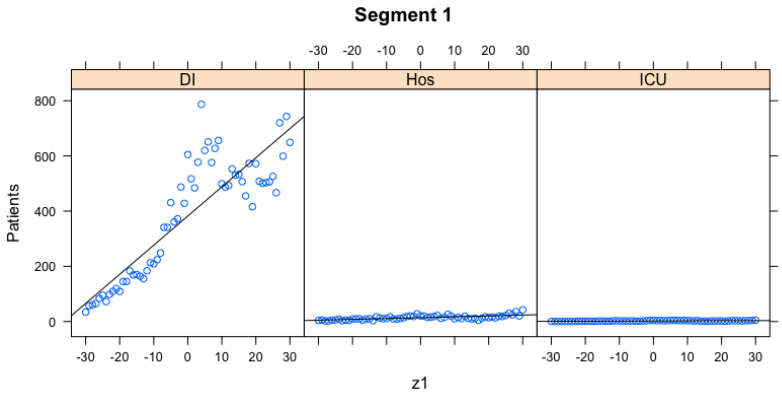

Figure 27: Centered model of Segment 1 of DI, Hos, and ICU

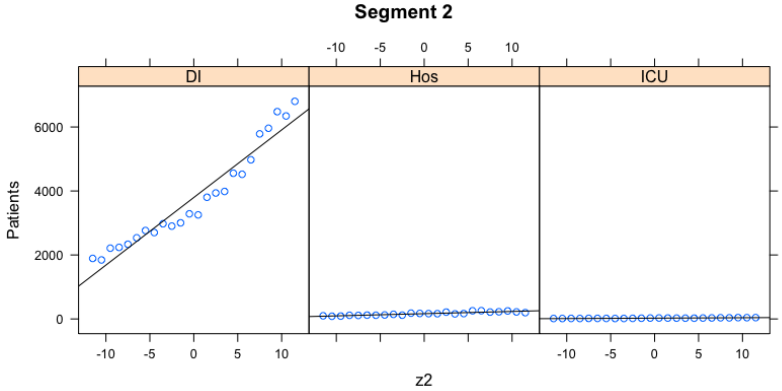

Figure 28: Centered model of Segment 2 of DI, Hos, and ICU

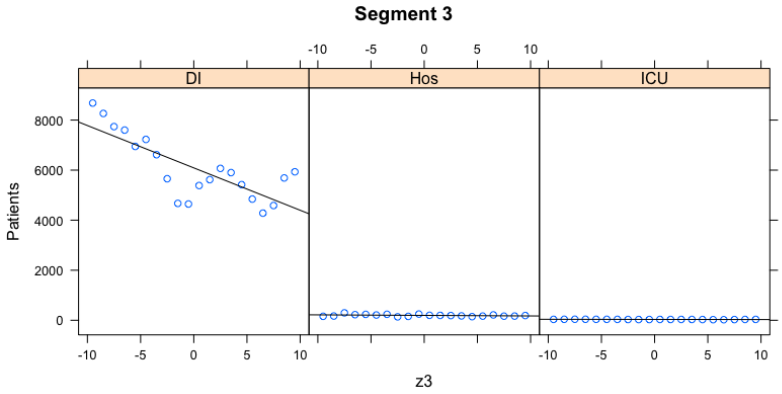

Figure 29: Centered model of Segment 3 of DI, Hos, and ICU

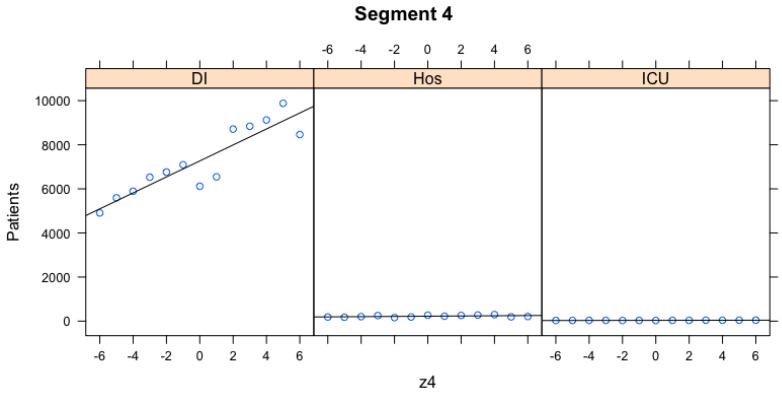

Figure 30: Centered model of Segment 4 of DI, Hos, and ICU

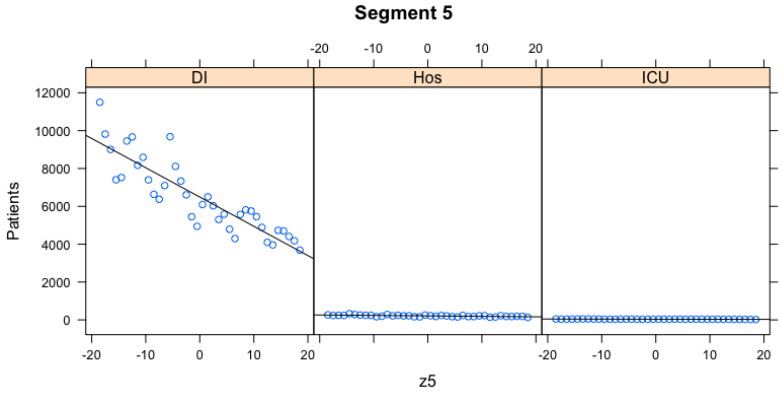

Figure 31: Centered model of Segment 5 of DI, Hos, and ICU

6.3.2 GLM form: Estimated Lines for Each Segment

This section is about the implementation of model in General Linear (GLM) form which allow different slopes, in the presence of interaction terms, that accommodate the three states DI, Hos, ICU.

The model in General Linear Model (GLM) form given by:

$$y_{ij} = \beta_0 + \tau_i + \beta_1 z_{ij} + \lambda_i z_{ij} + \epsilon_{ij} \qquad \begin{matrix} i = 1, ..., I \\ j = 1, ..., J \end{matrix} \tag{20}$$

with $\beta_0, \beta_0 + \tau_1, \beta_0 + \tau_2$ as the intercepts for state DI, state Hos, state ICU, respectively, while $\beta_1, \beta_1 + \lambda_1, \beta_1 + \lambda_2$ are the slopes for state DI, state Hos, state ICU, respectively. Index i denote the states and index j denote the generic days.

Further, one imposes an equal slope, i.e., imposing parallel lines for all the three states DI, Hos, ICU. This implies that the λ_i in equation (20) is set to be 0, i.e., setting both λ_1 and λ_2 to 0. This implementation is performed using R, as it is shown in Rasch et al. (2019) and Das and Mishra (2021). The F-test is then performed for all five segments with three states. Each F-test is being done per segment. After the F-test is performed, the p-values obtained are very close to zero from all five segments. These p-values obtained represent a highly significant p-value. Consequently, a close to zero p-value indicates that one can reject the null hypothesis, allowing one to conclude that the slopes are different between states. This step also concludes that it is not suitable to impose an equal slope.

The following are DI, Hos, and ICU in GLM form given in equation (21), (22), (23), respectively:

$$y_{ij} = \beta_0 + \beta_1 z_{ij} + \epsilon_{ij} \tag{21}$$

$$y_{ij} = (\beta_0 + \tau_1) + (\beta_1 + \lambda_1) z_{ij} + \epsilon_{ij} \tag{22}$$

$$y_{ij} = (\beta_0 + \tau_2) + (\beta_1 + \lambda_2) z_{ij} + \epsilon_{ij} \tag{23}$$

The following are the result for the line estimations:

Segment 1

Estimated line state DI: $\hat{y} = 382.18 + 10.56 z_1$

Estimated line state Hos: $\hat{y} = (382.18 - 368.23) + (10.56 - 10.25) z_1 = 13.95 + 0.31 z_1$

Estimated line state ICU: $\hat{y} = (382.18 - 380.13) + (10.56 - 10.51) z_1 = 2.05 + 0.05 z_1$

Segment 2

Estimated line state DI: $\hat{y} = 3795.54 + 211.29 z_2$

Estimated line state Hos: $\hat{y} = (3795.54 - 3630.71) + (211.29 - 204.40) z_2 = 164.83 + 6.89 z_2$

Estimated line state ICU: $\hat{y} = (3795.54 - 3770.63) + (211.29 - 210.05) z_2 = 24.91 + 1.24 z_2$

Segment 3

Estimated line state DI: $\hat{y} = 6088.45 - 169.35 z_3$

Estimated line state Hos: $\hat{y} = (6088.45 - 5894.90) + (-169.35 + 167.06)z_3 = 193.55 - 2.29z_3$
Estimated line state ICU: $\hat{y} = (6088.45 - 6055.40) + (-169.35 + 168.77)z_3 = 33.05 - 0.58z_3$

Segment 4
Estimated line state DI: $\hat{y} = 7266.23 + 362.42z_4$
Estimated line state Hos: $\hat{y} = (7266.23 - 7043.08) + (362.42 - 357.49)z_4 = 223.15 + 4.93z_4$
Estimated line state ICU: $\hat{y} = (7266.23 - 7230.54) + (362.42 - 361.24)z_4 = 35.69 + 1.18z_4$

Segment 5
Estimated line state DI: $\hat{y} = 6489.95 - 154.80z_5$
Estimated line state Hos: $\hat{y} = (6489.95 - 6279.05) + (-154.80 + 152.31)z_5 = 210.90 - 2.49z_5$
Estimated line state ICU: $\hat{y} = (6489.95 - 6453.42) + (-154.80 + 154.32)z_5 = 36.53 - 0.48z_5$

6.3.3 Retrieved Proportion

The line estimations for each state in each segment given in the section before are all obtained by utilizing a centered model (z_i). Generally, centering makes this value more interpretable because the expected value of y when centered (Points(z) is 0) represents the expected value of y when x is at its mean. In many cases, the intercept interpretation will be unreasonable or undesirable without some centering. It is also known that centering influences main effects in the presence of an interaction term. In the centered case, β_i is the main effect when the predictor variable x_i is equal to its mean. Hence, by centering, the interpretation of β_i remains the same when interactions are added.

Table 4: Proportion from Segment 1

	Points (z)						
	-30	-20	-10	0	10	20	30
Proportion DI-Hos	0.0711	0.0453	0.0392	0.0365	0.0349	0.0339	0.0333
Proportion Hos-ICU	0.1183	0.1355	0.1428	0.1469	0.1496	0.1514	0.1527

Table 5: Proportion from Segment 2

	Points (z)				
	-10	-5	0	5	10
Proportion DI-Hos	0.0570	0.0476	0.0434	0.0411	0.0395
Proportion Hos-ICU	0.1304	0.1435	0.1511	0.1561	0.1596

Table 6: Proportion from Segment 3

	Points (z)				
	-10	-5	0	5	10
Proportion DI-Hos	0.0278	0.0295	0.0318	0.0347	0.0388
Proportion Hos-ICU	0.1795	0.1754	0.1707	0.1656	0.1597

Table 7: Proportion from Segment 4

	Points (z)						
	-6	-4	-2	0	2	4	6
Proportion DI-Hos	0.0380	0.0349	0.0326	0.0307	0.0291	0.0279	0.0268
Proportion Hos-ICU	0.1478	0.1522	0.1563	0.1599	0.1633	0.1664	0.1692

Table 8: Proportion from Segment 5

	Points (z)				
	-20	-10	0	10	20
Proportion DI-Hos	0.0272	0.0293	0.0325	0.0376	0.0475
Proportion Hos-ICU	0.1769	0.1753	0.1732	0.1706	0.1672

Table 4 until Table 8 show the proportion of people that move from state DI to state Hos as well as the proportion of people that move from state Hos to state ICU over 5 or 7 data points. Instead of averaging the proportion over 5 or 7 points, one decided to choose the median from every segment. The reason is that any points above or below the median resulted in an unreliable proportion as it will be closer to the kink point to other segments.

Therefore, the proportion of people that move from state DI to state Hos is the average of the (red) median across the five segments, which is 0.035, and the proportion of people that move from state Hos to state ICU is the average of the (blue) median across five segments which is 0.160. Thus, approximately 3.5% of people are Detected Infected and will have to go to a hospital bed within approximately one week. In comparison, approximately 16% of people that are already in a hospital bed need to be placed in an intensive care unit within approximately one week. These proportions will be used for the next section of the implementation of Markov chain in order to see the evolution of patients over time.

6.4 Implementation of Markov Chains

In this section, the transition probability matrix is introduced first. Second, Markov chain simulations are performed to see how the chain evolves over time. The implementation of Markov chain simulations is performed using python 3.7.

6.4.1 Transition Probability Matrix with Five States

First, one has to construct the transition probability matrix. There will be five states: state DI, state Hos, state ICU, state DD, and state Recover (Rec). As obtained in section 6.3.3, the transition probability of detected infected people to a hospital bed (state DI to state Hos) is 0.035. Furthermore, transition probability of hospital bed to intensive care unit (state Hos to state ICU) is 0.160.

As mentioned in section 6.2.4 about an attempt of implementing SLS method on Dead (DD) data, there are two primary reasons as to why including DD data on the estimations by utilizing SLS method and least square method would lead to problems. However, the DD data is still necessary to this research. The existed paper, *NICE: COVID-19 op de Nederlandse Intensive Cares*, has complemented this research. NICE (2021) dynamics has provided this research the missing link of any transition that happens from state ICU, which are both state ICU to state DD and state ICU to state Hos. One assumes that from ICU is not possible to go straight to state DI.

Suppose that whenever the process is in state i, there is a probability p_{ij} that it will be in state j with $i, j \in \mathcal{S}$. Let $\mathcal{S}$ denoted the state space with $\mathcal{S} = \{DI, Hos, ICU, DD, Rec\}$. The following matrix P given in (24) shows the transition probability.

$$P = \begin{array}{c} \begin{matrix} \text{DI} & \text{Hos} & \text{ICU} & \text{DD} & \text{Rec} \end{matrix} \\ \begin{bmatrix} 0 & 0.035 & 0 & 0 & 0.965 \\ 0 & 0 & 0.160 & 0 & 0.840 \\ 0 & 0.705 & 0 & 0.268 & 0.027 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{matrix} \text{DI} \\ \text{Hos} \\ \text{ICU} \\ \text{DD} \\ \text{Rec} \end{matrix} \tag{24}$$

From the matrix P, the first row presents any transition that happens from state DI. Once the people are detected infected, they will never stay detected infected. This statement implies that there is a zero probability that state DI will stay in state DI. Detected Infected people would either in a worse condition than ended up needing medical support that can only be provided in a hospital bed, or they are recovering but not necessarily in a hospital bed. Hence, it is known that state DI will make a transition either to state Hos or state Rec. This statement implies there is also no chance of detected infected people that has to go to the intensive care unit straight away. Also, another zero chance of detected infected people is said to be dead right away. Thus, transition probability from state DI to state DI, from state DI to state ICU, and from state DI to DD are zero.

After thorough estimations being done in the last two sections, one can retrieve essential information about the expected proportion of people moving from

state DI to state Hos. Subsequently, there is an expected probability of 0.035 that detected infected people would have to go to a hospital bed to get medical assistance needed, leaving a 0.965 chance that detected infected people would end up independently recovering from SARS-CoV-2 outside the hospital.

The transitions that happen from state Hos are presented in the second row of matrix P. There are also only two possibilities that the people in state Hos will transition to. Again, when the condition is getting worse, the people in the hospital bed will need intensive medical care, which can be provided in an Intensive Care Unit. From an estimation in a previous section, one knows that the probability of ending up in an Intensive Care Unit from a hospital bed is 0.160. Hence, the transition probability from state Hos to state ICU is 0.160. There is also a chance that people in a hospital bed will end up recovering without ICU needed. Hence, there is a transition from state Hos to state Rec which in this case is 0.840.

Meanwhile, there is no transition from state Hos to state DI because one assumes that once become detected infected and become hospitalized, the person can only get worse that is ending up in ICU or recovering independently. There is also the assumption of not becoming infected the second time. Also, there is no transition from state Hos to state Hos since the people will not stay hospitalized forever. They will somehow move to either state ICU or state Rec. The last one from the second row is the transition from state Hos to state DD. This transition has a zero probability. In the Netherlands, the intensive care unit (ICU) bed has never been full, so if the people in hospital bed somehow become critical, they at least will get some intensive medical care first before passing away. This fact is the empirical reason why the transition from state Hos to state DD is zero. In other words, there is no way that state Hos transition to state DD right away since it has to transit to state ICU first.

The third row of matrix P presents the transition probability that happens from state ICU. State ICU will never go to state DI with the same assumption that there is no second infection. State ICU also will never stay in state ICU. Thus, the transition probability of state ICU to state DI and state ICU to state ICU are both zero.

However, according to NICE (2021), there is a chance that the people in the Intensive Care Unit will be dismissed to a regular hospital bed. This statement implies there is a transition probability from state ICU to state Hos, which is 0.705. There is also a chance that people in the Intensive Care Unit will end up dead with a probability of 0.268. Hence, leaving the probability of recovering directly after being in an Intensive Care Unit is 0.027.

The last two states in matrix P are state DD and state Recover. These are presented in the fourth and fifth rows, respectively. The transition that can happen is only state DD to state DD as once passed away then the people will remain so. As to state Rec will also stay in state Rec with probability 1. It is assumed that once recovered then the people will stay recovered. Thus, the transition probability from state DD to state DD and state Rec to state Rec are both one. Leaving no possibility of transitioning to other states than itself.

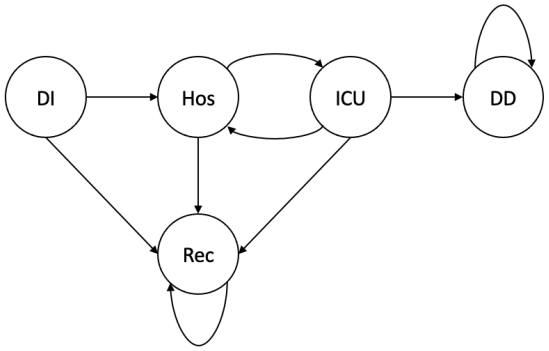

Figure 32: State Transitions Visualization

The visualization of how the state transitions is shown in Figure 32. It can be seen that regardless of where State DI goes, it will never return. Therefore, state DI is called a transient state. State Hos and state ICU communicate with each other. There are two absorbing states: state DD and state Rec as once the process gets there, the process will never leave this state. Therefore, the Markov chain in this research is then a reducible Markov chain.

6.4.2 Markov Chain Simulations

This section discussed the evolution of Markov chain in this research over time. The main question is: what is the probability that m steps from now, the process is in state j given that it is currently in state i?

The question given above is about n-step transition probabilities. In section 5.2, the n-step transition probabilities p_{ij}^n provides Markov chain performance characteristics. The Chapman-Kolmogorov equation states that the n-step transition probabilities $P^{(n)}$ can be obtained by multiplying the one-step transition probability matrix P by itself n times. Therefore, to understand the evolution of the Markov chain of this research over time, one can multiply the matrix P given in (24).

Also, notice that the chain in this research is a reducible Markov chain which means one of the three conditions to have a limiting distribution is not fulfilled. Not having a limiting distribution is not necessarily mean also not having a stationary distribution. Thus, there is a chance of having stationary distribution, which is beneficial to evaluate the long-run proportion. Hence, Markov chain simulation is performed to give more clarity about this matter.

Figure 33 gives the time scheme of n-steps transition. Additionally, Table 9 gives the result of the probabilities of being in states DI, Hos, ICU, DD, Rec, respectively, given the initial state is DI, n-steps from now. The number of transitions is shown as nrTransitions, i.e., n-steps. Each transition happens in approximately two weeks.

Thus, the first row implies that, given initial state DI, 3.57% of detected infected people must be in a hospital bed in approximately two weeks. Meanwhile, the rest of 96.43% will be recovering but not necessarily in a hospital bed. After four weeks, given state DI as the initial state, there is a 0.54% chance that they will be in the Intensive Care Unit while 99.56% are recovering.

Notice that there is a pattern obtained from this n-step transitions. It switches between having to go to a Hospital bed or ICU given the initial state of detected infected in the first seven steps. In other words, for the first fourteen weeks, detected infected people would still have a chance to go to a hospital bed or to an ICU unit.

As the number of steps grows, the probabilities then converge to 0.0017 and 0.9983 for being in state DD and state Rec, respectively. This statement also means that detected infected people would either pass away with a 0.17% chance or recover with a 99.83% chance starting from week sixteenth.

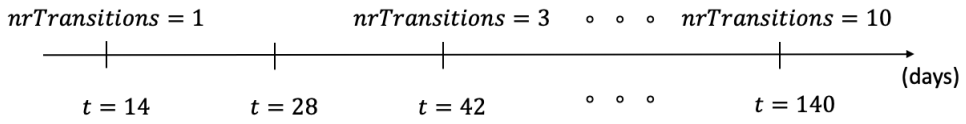

Figure 33: Time scheme of n-steps transitions

Table 9: The probabilities of being in states DI, Hos, ICU, DD, Rec, given state DI as initial state

nrTransitions	DI	Hos	ICU	DD	Rec
1	0	0.03573	0	0	0.96427
2	0	0	0.00541	0	0.99459
3	0	0.00371	0	0.0015	0.99479
4	0	0	0.00046	0.00163	0.99791
5	0	0.00046	0	0.00202	0.99752
6	0	0	0.00004	0.00194	0.99802
7	0	0.00005	0	0.00165	0.9983
8	0	0	0	0.00179	0.99821
9	0	0	0	0.00188	0.99812
10	0	0	0	0.00169	0.99831

The matrix $P^{(\infty)}$ in (25) presents the result of the stationary probabilities of being in each of the states DI, Hos, ICU, DD, Rec as $n \to \infty$. The first row of matrix $P^{(\infty)}$ implies that in the long run, there is a 99.8% chance of recovering given detected infected as the initial state. Meanwhile, there is a 95.17% chance of recovering given hospitalized as the initial state. Also, given the current condition of being in an ICU bed, there is a 69.79% chance of recovering. The last two rows showed that when the people passed away, it will remain so, and the same goes for people who already recovered.

$$P^{(\infty)} = \begin{array}{c} \begin{matrix} \text{DI} & \text{Hos} & \text{ICU} & \text{DD} & \text{Rec} \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 0.00169 & 0.99831 \\ 0 & 0 & 0 & 0.04833 & 0.95167 \\ 0 & 0 & 0 & 0.30207 & 0.69793 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{matrix} \text{DI} \\ \text{Hos} \\ \text{ICU} \\ \text{DD} \\ \text{Rec} \end{matrix} \tag{25}$$

6.5 An Analysis of SARS-CoV-2 in the Netherlands on Early Pandemic Days

This section is about getting to look closer at the condition of SARS-CoV-2 in the Netherlands from early March 2020 until mid-May 2020. In section 6.4.2, the probabilities of being in a certain state have been provided as well as the long-run proportion. This information then can be used in this chapter.

For example, given the first state 100,000 people detected infected, what will be the proportion of people in each state after two weeks?

One first define that the one-step transition period is equal to approximately two weeks. In order to answer this question, one needs to investigate the probability of being in certain states when the nrTransitions is equal to 1. There will be around 3,573 people ending up in hospital beds in approximately two weeks and around 96,427 people making an independent recovery, given 100,000 people were detected infected at time zero. However, this long-run proportion only gives the information of what will happen discretely. It only gives an idea of what will happen when the first wave hit after certain weeks. Subsequently, gives an idea of what will happen after the second wave hit after certain weeks and so on.

Next, one wanted to investigate further the SARS-CoV-2 data that are provided in *rijksoverheid.nl*. The visualization of this data are given in Figure 5 until Figure 8. Notice that from Figure 5 (DI data visualization), the early March 2020 until mid-May 2020 has a relatively flatter pattern compared to Hos, ICU, and DD data visualization. Thus, one wanted to reconstruct the detected infected people data in early pandemic days employing the transition probabilities from previous section.

The exact period that one is interested in observing is March 4th, 2020 until May 12th, 2020. This period has 70 days in total. One can reconstruct this period based on the transition probabilities, Hos data, and ICU data. This reconstruction aims to observe how the data in state DI would actually look like, given the same fashion of collecting the data with the period of July 2020 until January 2021, where a large scale of testing is implemented.

First, reconstruct the number of DI people given the Hos data and the transition probability from DI to Hos. Second, given the ICU data and the transition probabilities of Hos to ICU and DI to Hos, reconstruct the number of DI people. The latter compares the initial DI from the online dashboard with DI reconstruction from Hos data and DI reconstruction from ICU data. Figure 34 shows the visualization of the comparison. The blue line presents the initial DI data presented from RIVM, the orange line is DI data reconstruction from Hos data, and the grey line is DI data reconstruction from ICU data.

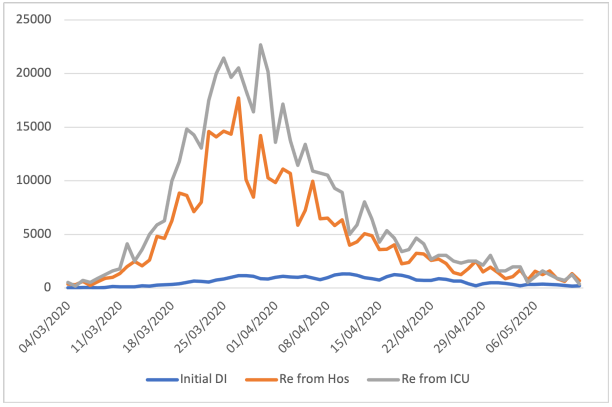

Figure 34: Comparison of Reconstructed DI data

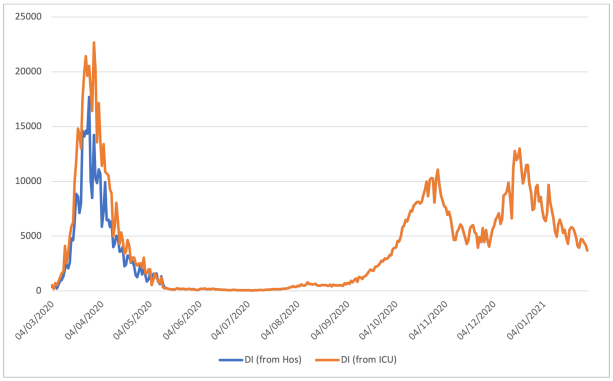

Figure 35: Reconstructed the first 70 days of DI placed back in the timeline

The reconstruction of these 70 days being placed back to the rest of the timeline can be seen in Figure 35. The blue line presents the first 70 days of reconstruction of DI from Hos data being placed back in the timeline, and the orange line presents the reconstruction from ICU data. The DI data supposedly has way larger numbers on the first 70 days if the same large scale of testing was also implemented in the early days of the pandemic.

Additionally, from the first 14 days, one estimates reproduction factor (R) by means of the following formula:

$$S_n = x_0 \left(\frac{1 - R^n}{1 - R} \right) \tag{26}$$

From (26), x_o is the first observation and S_n is the summation of the first n observations. The reproduction factor R is the number that shows how fast the virus is spreading. This factor shows how many people are on average infected by someone else who is already infected with the SARS-CoV-2 virus. This way, one can observe the number of transmissions that could be constant when the reproduction factor is 1 and declines as the reproduction factor is less than 1 but increases when the factor is above 1. The visualization of the reproduction with certain factors is shown in Figure 36.

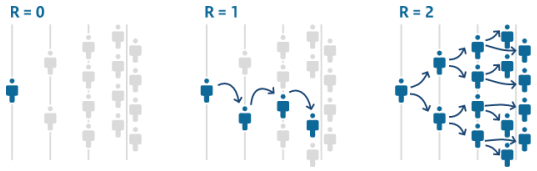

Figure 36: Visualization of Reproduction Factor (Source: RIVM)

On the average of 14 days, the reproduction factor is given in Table 10. In this table, the R of the DI data from the dashboard is approximately 1.65. While from the reconstruction of the number of DI people given Hos data leads to reproduction factor of approximately 2.06. The last one is the reconstruction of the number of DI people given ICU data gives the information of reproduction factor, which is around 2.13.

To sum up, to obtain this reproduction factor, one used the first 14 days, from March 4th until March 18th. The first reason as to why these 14 days are chosen to calculate R is that there is no governmental intervention at all during this period. After R is obtained, so one can roll the reconstruction based on the reproduction factor R to see how it could be if the government did not interfere over time. The second reason is that to make this calculation period is consistent with the transition period since all of the transition probabilities are adjusted to two weeks.

Table 10: Reproduction Factor (R)

Initial DI	From Hos	From ICU
1.65	2.06	2.13

6.6 Reconstruction of March 4th 2020 until May 12th 2020

As reproduction factor R is obtained from these first 14 days, now R is used to reconstruct the whole 70 days of DI people. Given the transition probabilities in the matrix (24), the following table shows how many people arrived in state DI, Hos, ICU in every two weeks transition.

Table 11: Transition Rollout of Reconstructed Data from Hos ($R = 2.06$)

nrTrans	Days Sum	DI	Hos	ICU
0	S_{14}	$23.9 \cdot 10^3$	839	134
1	$S_{28} - S_{14}$	$611 \cdot 10^6$	$21.4 \cdot 10^6$	$3.4 \cdot 10^6$
2	$S_{42} - S_{28}$	$15.6 \cdot 10^{12}$	$546 \cdot 10^9$	$87.4 \cdot 10^9$
3	$S_{56} - S_{42}$	$398.2 \cdot 10^{15}$	$13.9 \cdot 10^{15}$	$2.2 \cdot 10^{15}$
4	$S_{70} - S_{56}$	$10 \cdot 10^{21}$	$355 \cdot 10^{18}$	$56.9 \cdot 10^{18}$

Table 12: Transition Rollout of Reconstructed Data from ICU ($R = 2.13$)

nrTrans	Days Sum	DI	Hos	ICU
0	S_{14}	$34.8 \cdot 10^3$	$1.2 \cdot 10^3$	195
1	$S_{28} - S_{14}$	$1.4 \cdot 10^9$	$47.9 \cdot 10^6$	$7.6 \cdot 10^6$
2	$S_{42} - S_{28}$	$53.7 \cdot 10^{12}$	$1.8 \cdot 10^{12}$	$301 \cdot 10^9$
3	$S_{56} - S_{42}$	$2.1 \cdot 10^{18}$	$73.9 \cdot 10^{15}$	$11.8 \cdot 10^{15}$
4	$S_{70} - S_{56}$	$83 \cdot 10^{21}$	$2.9 \cdot 10^{21}$	$465.3 \cdot 10^{18}$

Table 11 and Table 12 show an overview of what would have happened if the government decided not to interfere with the spreading process of SARS-CoV-2 in the Netherlands during these early pandemic days. Also, the calculation and estimation are being done in a similar fashion of data collection from the online dashboard (period July 2020 until January 2021).

Table 11 shows the rollout of reconstructed data from Hos. For instance, given the start date is March 4th, the first transition (nrTrans = 1) suggest that around 611 million are detected infected in March 31st, which will bring around 21.4 million people out of them needed to have medical assistance in hospital bed around two weeks after (April 14th). Leaving around 3.4 million will eventually end up in ICU bed on April 28th. It happens according to the transition probability of people going from DI to Hos and the transition probability of people going from Hos to ICU, which are 0.035 and 0.16 in the matrix (24) as well as corresponding reproduction factor R of 2.06. This transition implies that in less than four weeks after March 4th, the whole population of the Netherlands could be infected if the government did not make any restrictions.

While according to Table 12 shows that on March 31st, there is around 1.4 billion people are detected infected, and around 47.9 million people out of them will need to be in a hospital bed in around two weeks after, while 7.6 million will end up in ICU bed in around four weeks after March 31st. This estimation is according to the exact transition probabilities and corresponding reproduction factor R of 2.13. The estimation seems to increase drastically given an increase in reproduction factor R, though it is only increased by 0.07.

7 Summary, Conclusions, and Recommendations

This thesis started with a review on the online dashboard *rijksoverheid.nl* which provided four types of SARS-CoV-2 data in the Netherlands. These data are Detected Infected (DI), Hospitalized (Hos), Intensive Care Unit (ICU), and Dead (DD).

Segmented Least Squares (SLS) is implemented to DI data first. The period chosen for this estimation is from July 7th, 2020, until January 31st, 2021. The optimal segmentation obtained is five segments. Then, SLS is applied to Hos and ICU data. It turned out SLS also gives desirable patterns in five segments. So, SLS resulted in five optimal segmentation for all three types/states (DI, Hos, and ICU data). SLS implementation on DD data is not being taken into account since the result would not be reliable. Subsequently, the assumption of window shift between states is validated statistically. It is proven that the window of detected infected people to shift to hospital beds is on average seven days. In comparison, the patients in hospital beds will, on average, shift to ICU in two days.

Next to that, regression in the presence of a qualitative factor was performed. The F-statistic is performed on the model in general linear form. It turned out that the lines are not parallel; therefore, imposing a parallel line is not suitable. The lines estimation allowing different slopes then obtained. The proportion of DI to Hos and the proportion of Hos to ICU are obtained by utilizing the centered model of the estimated lines. These proportions are used to build the transition probability matrix P with five states: DI, Hos, ICU, DD, Rec. The transition probability from ICU to Hos and ICU to DD are given from NICE (2021). With P matrix, Markov chain evolution can then be observed.

In addition, an analysis of SARS-CoV-2 on early days is performed. It turned out that the DI data shown in the online dashboard *rijksoverheid.nl* is not being reasonable since it supposedly has the highest peak around the end of March 2020. The reasonable reproduction factors are also obtained. Lastly, the reconstruction of the early pandemic day is performed. This reconstruction shows how fast the spreading of SARS-CoV-2 would have been if the government decided not to interfere with any measurements. It is shown that around early April 2020, the whole population of the Netherlands could have been infected if there were no restrictions. This could happen since the reproduction factor was high at the start of the pandemic.

In conclusion, dynamic programming on multi-way choices has validated that there are window shifts between states. Then, the centered model in general linear form resulted in proportions of one state to move to another. These proportions then become transition probabilities in Markov Chain. The transition probabilities matrix rollout showed that the actual condition on early pandemic days was much worse than shown in the online dashboard. It would have taken only four more weeks for the virus to spread quickly throughout the whole population of the Netherlands if there was no governmental interference. Following infections, hospitalization and intensive care need would have been huge, leading to disastrous consequences for the Dutch society.

A recommendation for future work would be to develop a different method for calculating the day shift mismatch of all possible combinations. Currently, the calculation is done manually throughout 72 possible combinations with five

segments and three states in each combination.

A second recommendation would be to observe further the possibility of becoming infected the second time. In this thesis, one assumes that being infected the second time is not possible. However, given the current delta mutation globally, infected people can get infected the second time after being recovered.

Next to that, another idea of elaboration is to further observe the kink from the patterns with respected governmental interference during that kink period. This elaboration then could be studied further with Markov Decision Processes instead.

References

Bellman, R. and Roth, R. (1969). Curve fitting by segmented straight lines. *Journal of the American Statistical Association*, 64(327):1079–1084.

CDC (2020). Coronavirus disease 2019 (covid-19): people who need to take extra precautions. `https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/index.html`. Accessed: 2021-03-08.

Das, I. and Mishra, A. (2021). *Trends of Data Science and Applications*, volume 954 of *Studies in Computational Intelligence*. Springer, Singapore.

Garg, S., Kim, L., and Whitaker, M. (2020). Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019. *Morbidity and Mortality Weekly Report*, 69:458–464.

Kim, H.-J., Yu, B., and Feuer, E. J. (2008). Inference in segmented line regression: a simulation study. *Journal of Statistical Computation and Simulation*, 78(11):1087–1103.

Kleinberg, J. and Tardos, E. (2005). *Algorithm Design*. Pearson Addison Wesley, Boston, 1st edition.

Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2005). *Applied Linear Statistical Models*. McGraw-Hill Irwin, 5th edition.

Mankowski, M. and Moshkov, M. (2021). *Dynamic Programming Multi-Objective Combinatorial Optimization*, volume 331 of *Studies in Systems, Decision and Control*. Springer, Cham.

McZgee, V. E. and Carleton, W. T. (1970). Piecewise regression. *Journal of the American Statistical Association*, 65(331):1109–1124.

NICE (2021). Covid-19 op de nederlandse intensive cares. `https://www.stichting-nice.nl/COVID_rapport.pdf`. Accessed: 2021-05-04.

Pinsky, M. A. and Karlin, S. (2011). *Introduction to Stochastic Modelling*. Academic Press, 4th edition.

Rasch, D., Verdooren, R., and Pilz, J. (2019). *Applied Statistics: Theory and Problem Solutions with R*. WILEY.

RIVM (2020). Coronavirus disease covid-19: Incubation period. `https://www.rivm.nl/en/coronavirus-covid-19/coronavirus-disease-covid-19`. Accessed: 2021-03-08.

Ross, S. M. (2014). *Introduction to Probability Models*. Academic Press, 11th edition.

Rutherford, A. (2001). *Introducing Anova and Ancova : A GLM Approach*. SAGE Publications.

WHO (2020). Coronavirus: Overview. `https://www.who.int/health-topics/coronavirus#tab=tab_1`. Accessed: 2021-03-08.

Appendices

A Appendix: Day shifts Tables

A.1 DI-Hos shift on Day 7th

Table 13: DI-Hos: Day 7th, Hos-ICU: Day 1st

	DI	Hos	ICU
Segment 1	1-74	1-61	1-69
Segment 2	75-116	62-112	70-99
Segment 3	118-147	114-137	101-138
Segment 4	149-161	139-168	140-171
Segment 5	163-209	170-209	173-209

Table 14: DI-Hos: Day 7th, Hos-ICU: Day 2nd

	DI	Hos	ICU
Segment 1	1-74	1-61	1-68
Segment 2	75-116	62-112	69-98
Segment 3	118-147	114-137	100-137
Segment 4	149-161	139-168	139-170
Segment 5	163-209	170-209	172-209

Table 15: DI-Hos: Day 7th, Hos-ICU: Day 3rd

	DI	Hos	ICU
Segment 1	1-74	1-61	1-67
Segment 2	75-116	62-112	68-97
Segment 3	118-147	114-137	99-136
Segment 4	149-161	139-168	138-169
Segment 5	163-209	170-209	171-209

Table 16: DI-Hos: Day 7th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-61	1-65
Segment 2	75-116	62-112	66-95
Segment 3	118-147	114-137	97-134
Segment 4	149-161	139-168	136-167
Segment 5	163-209	170-209	169-209

Table 17: DI-Hos: Day 7th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-61	1-64
Segment 2	75-116	62-112	65-94
Segment 3	118-147	114-137	96-133
Segment 4	149-161	139-168	135-166
Segment 5	163-209	170-209	168-209

Table 18: DI-Hos: Day 7th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-61	1-63
Segment 2	75-116	62-112	64-93
Segment 3	118-147	114-137	95-132
Segment 4	149-161	139-168	134-165
Segment 5	163-209	170-209	167-209

Table 19: DI-Hos: Day 7th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-61	1-62
Segment 2	75-116	62-112	63-92
Segment 3	118-147	114-137	94-131
Segment 4	149-161	139-168	133-164
Segment 5	163-209	170-209	166-209

Table 20: DI-Hos: Day 7th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-61	1-61
Segment 2	75-116	62-112	62-91
Segment 3	118-147	114-137	93-130
Segment 4	149-161	139-168	132-163
Segment 5	163-209	170-209	165-209

A.2 DI-Hos shift on Day 8th

Table 21: DI-Hos: Day 8th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-60	1-65
Segment 2	75-116	61-111	66-95
Segment 3	118-147	113-136	97-134
Segment 4	149-161	138-167	136-167
Segment 5	163-209	169-209	169-209

Table 22: DI-Hos: Day 8th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-60	1-64
Segment 2	75-116	61-111	65-94
Segment 3	118-147	113-136	96-133
Segment 4	149-161	138-167	135-166
Segment 5	163-209	169-209	168-209

Table 23: DI-Hos: Day 8th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-60	1-63
Segment 2	75-116	61-111	64-93
Segment 3	118-147	113-136	95-132
Segment 4	149-161	138-167	134-165
Segment 5	163-209	169-209	167-209

Table 24: DI-Hos: Day 8th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-60	1-62
Segment 2	75-116	61-111	63-92
Segment 3	118-147	113-136	94-131
Segment 4	149-161	138-167	133-164
Segment 5	163-209	169-209	166-209

Table 25: DI-Hos: Day 8th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-60	1-61
Segment 2	75-116	61-111	62-91
Segment 3	118-147	113-136	93-130
Segment 4	149-161	138-167	132-163
Segment 5	163-209	169-209	165-209

Table 26: DI-Hos: Day 8th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-60	1-60
Segment 2	75-116	61-111	61-90
Segment 3	118-147	113-136	92-129
Segment 4	149-161	138-167	131-162
Segment 5	163-209	169-209	164-209

A.3 DI-Hos shift on Day 9th

Table 27: DI-Hos: Day 9th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-59	1-64
Segment 2	75-116	60-110	65-94
Segment 3	118-147	112-135	96-133
Segment 4	149-161	137-166	135-166
Segment 5	163-209	168-209	168-209

Table 28: DI-Hos: Day 9th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-59	1-63
Segment 2	75-116	60-110	64-93
Segment 3	118-147	112-135	95-132
Segment 4	149-161	137-166	134-165
Segment 5	163-209	168-209	167-209

Table 29: DI-Hos: Day 9th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-59	1-62
Segment 2	75-116	60-110	63-92
Segment 3	118-147	112-135	94-131
Segment 4	149-161	137-166	133-164
Segment 5	163-209	168-209	166-209

Table 30: DI-Hos: Day 9th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-59	1-61
Segment 2	75-116	60-110	62-91
Segment 3	118-147	112-135	93-130
Segment 4	149-161	137-166	132-163
Segment 5	163-209	168-209	165-209

Table 31: DI-Hos: Day 9th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-59	1-60
Segment 2	75-116	60-110	61-90
Segment 3	118-147	112-135	92-129
Segment 4	149-161	137-166	131-162
Segment 5	163-209	168-209	164-209

Table 32: DI-Hos: Day 9th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-59	1-59
Segment 2	75-116	60-110	60-89
Segment 3	118-147	112-135	91-128
Segment 4	149-161	137-166	130-161
Segment 5	163-209	168-209	163-209

A.4 DI-Hos shift on Day 10th

Table 33: DI-Hos: Day 10th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-58	1-63
Segment 2	75-116	69-109	64-93
Segment 3	118-147	111-134	95-132
Segment 4	149-161	136-165	134-165
Segment 5	163-209	167-209	167-209

Table 34: DI-Hos: Day 10th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-58	1-62
Segment 2	75-116	69-109	63-92
Segment 3	118-147	111-134	94-131
Segment 4	149-161	136-165	133-164
Segment 5	163-209	167-209	166-209

Table 35: DI-Hos: Day 10th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-58	1-61
Segment 2	75-116	69-109	62-91
Segment 3	118-147	111-134	93-130
Segment 4	149-161	136-165	132-163
Segment 5	163-209	167-209	165-209

Table 36: DI-Hos: Day 10th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-58	1-60
Segment 2	75-116	69-109	61-90
Segment 3	118-147	111-134	92-129
Segment 4	149-161	136-165	131-162
Segment 5	163-209	167-209	164-209

Table 37: DI-Hos: Day 10th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-58	1-59
Segment 2	75-116	69-109	60-89
Segment 3	118-147	111-134	91-128
Segment 4	149-161	136-165	130-161
Segment 5	163-209	167-209	163-209

Table 38: DI-Hos: Day 10th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-58	1-58
Segment 2	75-116	69-109	59-88
Segment 3	118-147	111-134	90-127
Segment 4	149-161	136-165	129-160
Segment 5	163-209	167-209	162-209

A.5 DI-Hos shift on Day 11th

Table 39: DI-Hos: Day 11th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-57	1-62
Segment 2	75-116	58-108	63-92
Segment 3	118-147	110-133	94-131
Segment 4	149-161	135-164	133-164
Segment 5	163-209	166-209	166-209

Table 40: DI-Hos: Day 11th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-57	1-61
Segment 2	75-116	58-108	62-91
Segment 3	118-147	110-133	93-130
Segment 4	149-161	135-164	132-163
Segment 5	163-209	166-209	165-209

Table 41: DI-Hos: Day 11th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-57	1-60
Segment 2	75-116	58-108	61-90
Segment 3	118-147	110-133	92-129
Segment 4	149-161	135-164	131-162
Segment 5	163-209	166-209	164-209

Table 42: DI-Hos: Day 11th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-57	1-59
Segment 2	75-116	58-108	60-89
Segment 3	118-147	110-133	91-128
Segment 4	149-161	135-164	130-161
Segment 5	163-209	166-209	163-209

Table 43: DI-Hos: Day 11th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-57	1-58
Segment 2	75-116	58-108	59-88
Segment 3	118-147	110-133	90-127
Segment 4	149-161	135-164	129-160
Segment 5	163-209	166-209	162-209

Table 44: DI-Hos: Day 11th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-57	1-57
Segment 2	75-116	58-108	58-87
Segment 3	118-147	110-133	89-126
Segment 4	149-161	135-164	128-160
Segment 5	163-209	166-209	162-209

A.6 DI-Hos shift on Day 12th

Table 45: DI-Hos: Day 12th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-56	1-61
Segment 2	75-116	57-107	62-91
Segment 3	118-147	109-132	93-130
Segment 4	149-161	134-163	132-163
Segment 5	163-209	165-209	165-209

Table 46: DI-Hos: Day 12th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-56	1-60
Segment 2	75-116	57-107	61-90
Segment 3	118-147	109-132	92-129
Segment 4	149-161	134-163	131-162
Segment 5	163-209	165-209	164-209

Table 47: DI-Hos: Day 12th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-56	1-59
Segment 2	75-116	57-107	60-89
Segment 3	118-147	109-132	91-128
Segment 4	149-161	134-163	130-161
Segment 5	163-209	165-209	163-209

Table 48: DI-Hos: Day 12th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-56	1-58
Segment 2	75-116	57-107	59-88
Segment 3	118-147	109-132	90-127
Segment 4	149-161	134-163	129-160
Segment 5	163-209	165-209	162-209

Table 49: DI-Hos: Day 12th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-56	1-57
Segment 2	75-116	57-107	58-87
Segment 3	118-147	109-132	89-126
Segment 4	149-161	134-163	128-159
Segment 5	163-209	165-209	161-209

Table 50: DI-Hos: Day 12th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-56	1-56
Segment 2	75-116	57-107	57-86
Segment 3	118-147	109-132	88-125
Segment 4	149-161	134-163	127-159
Segment 5	163-209	165-209	161-209

A.7 DI-Hos shift on Day 13th

Table 51: DI-Hos: Day 13th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-55	1-60
Segment 2	75-116	56-106	61-90
Segment 3	118-147	108-131	92-129
Segment 4	149-161	133-162	131-162
Segment 5	163-209	164-209	164-209

Table 52: DI-Hos: Day 13th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-55	1-59
Segment 2	75-116	56-106	60-89
Segment 3	118-147	108-131	91-128
Segment 4	149-161	133-162	130-161
Segment 5	163-209	164-209	163-209

Table 53: DI-Hos: Day 13th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-55	1-58
Segment 2	75-116	56-106	59-88
Segment 3	118-147	108-131	90-127
Segment 4	149-161	133-162	129-160
Segment 5	163-209	164-209	162-209

Table 54: DI-Hos: Day 13th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-55	1-57
Segment 2	75-116	56-106	58-87
Segment 3	118-147	108-131	89-126
Segment 4	149-161	133-162	128-159
Segment 5	163-209	164-209	161-209

Table 55: DI-Hos: Day 13th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-55	1-56
Segment 2	75-116	56-106	57-86
Segment 3	118-147	108-131	88-125
Segment 4	149-161	133-162	127-158
Segment 5	163-209	164-209	160-209

Table 56: DI-Hos: Day 13th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-55	1-55
Segment 2	75-116	56-106	56-85
Segment 3	118-147	108-131	87-124
Segment 4	149-161	133-162	126-158
Segment 5	163-209	164-209	160-209

A.8 DI-Hos shift on Day 14th

Table 57: DI-Hos: Day 14th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-54	1-59
Segment 2	75-116	55-105	60-89
Segment 3	118-147	107-130	91-128
Segment 4	149-161	132-161	130-161
Segment 5	163-209	163-209	163-209

Table 58: DI-Hos: Day 14th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-54	1-58
Segment 2	75-116	55-105	59-88
Segment 3	118-147	107-130	90-127
Segment 4	149-161	132-161	129-160
Segment 5	163-209	163-209	162-209

Table 59: DI-Hos: Day 14th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-54	1-57
Segment 2	75-116	55-105	58-87
Segment 3	118-147	107-130	89-126
Segment 4	149-161	132-161	128-159
Segment 5	163-209	163-209	161-209

Table 60: DI-Hos: Day 14th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-54	1-56
Segment 2	75-116	55-105	57-86
Segment 3	118-147	107-130	88-125
Segment 4	149-161	132-161	127-158
Segment 5	163-209	163-209	160-209

Table 61: DI-Hos: Day 14th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-54	1-55
Segment 2	75-116	55-105	56-85
Segment 3	118-147	107-130	87-124
Segment 4	149-161	132-161	126-157
Segment 5	163-209	163-209	159-209

Table 62: DI-Hos: Day 14th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-54	1-54
Segment 2	75-116	55-105	55-84
Segment 3	118-147	107-130	86-123
Segment 4	149-161	132-161	125-157
Segment 5	163-209	163-209	159-209

A.9 DI-Hos shift on Day 15th

Table 63: DI-Hos: Day 15th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-53	1-58
Segment 2	75-116	54-104	59-88
Segment 3	118-147	106-129	90-127
Segment 4	149-161	131-160	129-160
Segment 5	163-209	162-209	162-209

Table 64: DI-Hos: Day 15th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-53	1-57
Segment 2	75-116	54-104	58-87
Segment 3	118-147	106-129	89-126
Segment 4	149-161	131-160	128-159
Segment 5	163-209	162-209	161-209

Table 65: DI-Hos: Day 15th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-53	1-56
Segment 2	75-116	54-104	57-86
Segment 3	118-147	106-129	88-125
Segment 4	149-161	131-160	127-158
Segment 5	163-209	162-209	160-209

Table 66: DI-Hos: Day 15th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-53	1-55
Segment 2	75-116	54-104	56-85
Segment 3	118-147	106-129	87-124
Segment 4	149-161	131-160	126-157
Segment 5	163-209	162-209	159-209

Table 67: DI-Hos: Day 15th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-53	1-54
Segment 2	75-116	54-104	55-84
Segment 3	118-147	106-129	86-123
Segment 4	149-161	131-160	125-156
Segment 5	163-209	162-209	158-209

Table 68: DI-Hos: Day 15th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-53	1-53
Segment 2	75-116	54-104	54-83
Segment 3	118-147	106-129	85-122
Segment 4	149-161	131-160	124-156
Segment 5	163-209	162-209	158-209

A.10 DI-Hos shift on Day 16th

Table 69: DI-Hos: Day 16th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-52	1-57
Segment 2	75-116	53-103	58-87
Segment 3	118-147	105-128	89-126
Segment 4	149-161	130-159	128-160
Segment 5	163-209	161-209	162-209

Table 70: DI-Hos: Day 16th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-52	1-56
Segment 2	75-116	53-103	57-86
Segment 3	118-147	105-128	88-125
Segment 4	149-161	130-159	127-159
Segment 5	163-209	161-209	161-209

Table 71: DI-Hos: Day 16th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-52	1-55
Segment 2	75-116	53-103	56-85
Segment 3	118-147	105-128	87-124
Segment 4	149-161	130-159	126-158
Segment 5	163-209	161-209	160-209

Table 72: DI-Hos: Day 16th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-52	1-54
Segment 2	75-116	53-103	55-84
Segment 3	118-147	105-128	86-123
Segment 4	149-161	130-159	125-157
Segment 5	163-209	161-209	159-209

Table 73: DI-Hos: Day 16th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-52	1-53
Segment 2	75-116	53-103	54-83
Segment 3	118-147	105-128	85-122
Segment 4	149-161	130-159	124-156
Segment 5	163-209	161-209	158-209

Table 74: DI-Hos: Day 16th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-52	1-52
Segment 2	75-116	53-103	53-82
Segment 3	118-147	105-128	84-121
Segment 4	149-161	130-159	123-155
Segment 5	163-209	161-209	157-209

A.11 DI-Hos shift on Day 17th

Table 75: DI-Hos: Day 17th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-51	1-56
Segment 2	75-116	52-102	57-86
Segment 3	118-147	104-127	88-125
Segment 4	149-161	129-158	127-159
Segment 5	163-209	160-209	161-209

Table 76: DI-Hos: Day 17th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-51	1-55
Segment 2	75-116	52-102	56-85
Segment 3	118-147	104-127	87-124
Segment 4	149-161	129-158	126-158
Segment 5	163-209	160-209	160-209

Table 77: DI-Hos: Day 17th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-51	1-54
Segment 2	75-116	52-102	55-84
Segment 3	118-147	104-127	86-123
Segment 4	149-161	129-158	125-157
Segment 5	163-209	160-209	159-209

Table 78: DI-Hos: Day 17th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-51	1-53
Segment 2	75-116	52-102	54-83
Segment 3	118-147	104-127	85-122
Segment 4	149-161	129-158	124-156
Segment 5	163-209	160-209	158-209

Table 79: DI-Hos: Day 17th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-51	1-52
Segment 2	75-116	52-102	53-82
Segment 3	118-147	104-127	84-121
Segment 4	149-161	129-158	123-155
Segment 5	163-209	160-209	157-209

Table 80: DI-Hos: Day 17th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-51	1-51
Segment 2	75-116	52-102	52-81
Segment 3	118-147	104-127	83-120
Segment 4	149-161	129-158	122-154
Segment 5	163-209	160-209	156-209

A.12 DI-Hos shift on Day 18th

Table 81: DI-Hos: Day 18th, Hos-ICU: Day 4th

	DI	Hos	ICU
Segment 1	1-74	1-50	1-55
Segment 2	75-116	51-101	56-85
Segment 3	118-147	103-126	87-124
Segment 4	149-161	128-157	126-158
Segment 5	163-209	159-209	160-209

Table 82: DI-Hos: Day 18th, Hos-ICU: Day 5th

	DI	Hos	ICU
Segment 1	1-74	1-50	1-54
Segment 2	75-116	51-101	55-84
Segment 3	118-147	103-126	86-123
Segment 4	149-161	128-157	125-157
Segment 5	163-209	159-209	159-209

Table 83: DI-Hos: Day 18th, Hos-ICU: Day 6th

	DI	Hos	ICU
Segment 1	1-74	1-50	1-53
Segment 2	75-116	51-101	54-83
Segment 3	118-147	103-126	85-122
Segment 4	149-161	128-157	124-156
Segment 5	163-209	159-209	158-209

Table 84: DI-Hos: Day 18th, Hos-ICU: Day 7th

	DI	Hos	ICU
Segment 1	1-74	1-50	1-52
Segment 2	75-116	51-101	53-82
Segment 3	118-147	103-126	84-121
Segment 4	149-161	128-157	123-155
Segment 5	163-209	159-209	157-209

Table 85: DI-Hos: Day 18th, Hos-ICU: Day 8th

	DI	Hos	ICU
Segment 1	1-74	1-50	1-51
Segment 2	75-116	51-101	52-81
Segment 3	118-147	103-126	83-120
Segment 4	149-161	128-157	122-154
Segment 5	163-209	159-209	156-209

Table 86: DI-Hos: Day 18th, Hos-ICU: Day 9th

	DI	Hos	ICU
Segment 1	1-74	1-50	1-50
Segment 2	75-116	51-101	51-80
Segment 3	118-147	103-126	82-119
Segment 4	149-161	128-157	121-153
Segment 5	163-209	159-209	155-209