Nishchay Pandya

Nishchaypandya1210@gmail.com

Abstract

The Markov process is used to simulate and analyse the types of benefits received by the citizens of an unemployment benefit program. Additionally, discrete-time Markov chain is used to model the transition of citizens between different states and to project the outcomes and future costs of two types of unemployment programs

Stochastic modelling project

Markov Process

Table of Contents

[Transition Probability Matrices 2](#_bookmark0)

[Current model 2](#_bookmark1)

[Updated model without intervention 2](#_bookmark2)

[. 2](#_bookmark3)

[Updated model with intervention 2](#_bookmark4)

[. 2](#_bookmark5)

[Annual total costs distribution for the current and updated model 3](#_bookmark6)

[Differences 3](#_bookmark7)

[Similarities 3](#_bookmark8)

[Expected long-run costs 4](#_bookmark9)

[Discussion 4](#_bookmark10)

[Advantages and disadvantages of both models 4](#_bookmark11)

[Monte Carlo simulation and normal distribution 4](#_bookmark12)

[Improvements 5](#_bookmark13)

[Assumptions, limitations and constraints 5](#_bookmark14)

[Conclusion 5](#_bookmark15)

[Appendix 6](#_bookmark16)

[. 6](#_bookmark17)

[. 7](#_bookmark18)

[. 7](#_bookmark19)

[Bibliography 7](#_bookmark19)

# Transition Probability Matrices

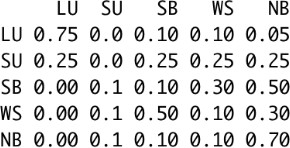
## Current model

To represent the current model as a discrete time Markov chain, a 4x4 matrix is constructed with the state space: S = {UB, SB, WS, NB}, where UB is unemployment benefit, SB is studying benefit, WS is wage subsidy, and NB is no benefit from the unemployment support program. Figure 1.2 is a visualisation depicting the state diagram of the current model, from which it can be deduced that all states communicate with each other, and hence it has only one communicating class. Therefore, this Markov chain is irreducible. Since this is a finite Markov chain, at least one state must be recurrent, and since there is only one communicating class, all states are recurrent. The period of state one is equal to one, as the probability of returning to state one with n step transitions for all n > 0 is positive, and one is the greatest common divisor of {1,2, 3, …., n}. Since all states communicate, the period of all states is equal to one, and hence this Markov chain is aperiodic and irreducible, which implies that it is ergodic. These properties are vital to analyse the business situation, as an ergodic chain will approach an equilibrium situation (limiting probability) when the number of transition steps are large (long run), which will have further applications in task 2.1 in calculating long run (constant) cost for the program(s).

***Figure*** 1.1

***Figure*** 1.2

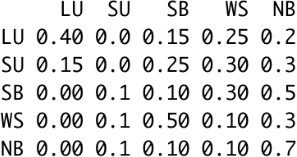
## Updated model without intervention



***Figure*** 1.3

To represent the updated model without intervention, a 5x5 transition probability matrix was constructed, with the additional state of LU referring to long term unemployment, and SU refers to short term unemployment. Therefore, the new state space is S = {LU, SU, WS, SB, NB}. To similar reasonings outlined for the current model, the Markov chain for the updated model without intervention is ergodic.

## Updated model with intervention

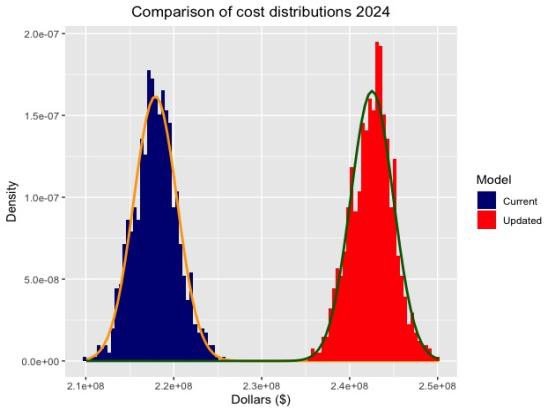
The transition probability matrix for the updated model with intervention is constituent of the same state space defined for the updated model without intervention. Similar to reasonings highlighted in the current model, the Markov chain for this model is ergodic, and will have further implications in task 2.1.

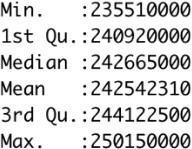
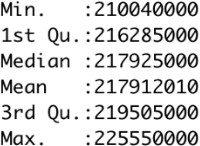
***Figure*** 1.4

# Annual total costs distribution for the current and updated model

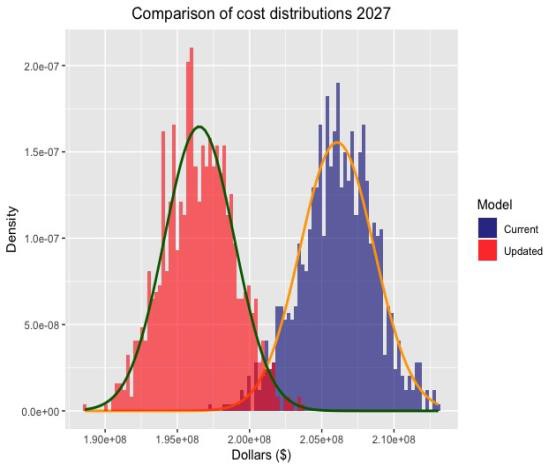
Using the discrete-time Markov chain transition matrices in task 1.1 and the given data of the citizens’ benefit program in 2021 and 2022, 1000 simulations of the transition of benefit programs were performed for the years 2023-2027 under the current, and both types of the updated model.

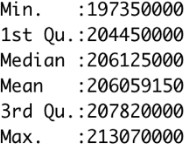
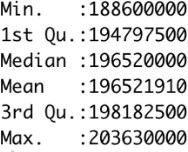
## Differences

In 2024, the mean value for the updated model’s cost is 11.30% greater than mean of the current model’s cost (for this particular simulation), and such a large difference can be explained by the 20,000,000 Happyland dollars intervention cost allocated to the updated model in 2024. The standard deviation for the current model’s cost is 2% greater than updated model’s standard deviation, which suggests slightly more uncertainty in expecting the cost for the current model.



***Figure*** 1.5

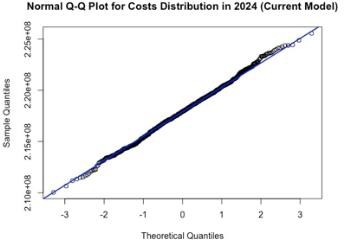
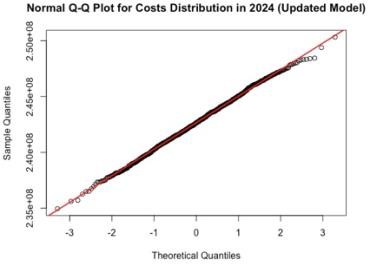
In 2027, the mean for the updated model cost is 4.63% less than the mean of the current model cost. Regarding the variable costs, the success of the intervention program has reduced the number of unemployed citizens, and therefore the cost to the government has decreased. Additionally, the fixed cost has been reduced to 1,000,000 Happyland dollars from 2025, which is significantly cheaper compared to the prior fixed intervention cost. Similar to 2024, the standard deviation for the current model’s cost is 5.85% higher than updated model’s standard deviation, indicating a greater dispersion of values for the current costs’ distribution.



***Figure*** 1.6

## Similarities

A similarity for both distributions, is that their mean and median are almost identical, which is a property of the normal distribution. All of the costs’ distributions follow a normal (Gaussian) distribution, which is depicted by the precise overlay of the normal density curve to their respective histograms. To investigate the conjecture that the annual total costs data for the current and updated model follow a normal distribution, a normal Q-Q plot was constructed for both costs, and as shown from both plots, the data set on both ends of the Q-Q plot do not deviate from the Q-Q line, suggesting that both costs series follow a normal distribution. A similar Q-Q plot can be obtained for the 2027 cost distributions, which will be included in the appendix.



***Figure*** 1.8

***Figure*** 1.7

Furthermore, the Shapiro-Wilk normality test was conducted using R studio, and the p-value for the cost series were greater than 0.05, and hence we cannot reject the null hypothesis of the data following a normal distribution.

# Expected long-run costs

The long run cost of both models further illuminates the difference in their effectiveness. As both models’ Markov chains are ergodic, the equilibrium situation was obtained by calculating the 100-step transition matrix for both Markov chains. The limiting probability property ensured that the distribution of citizens in each benefit program will remain constant and independent of the initial state, and the fixed distribution of citizens was multiplied with the vectorized payment plan, which results into the expected long run cost.

The expected long run cost for the current and updated model is: 205590062 and 194091787 Happyland dollars respectively. The annual long run cost for the current model is 5.56% more expensive than the updated model with intervention.

***Figure*** 1.9

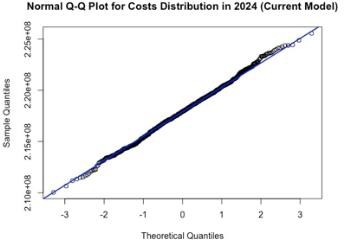
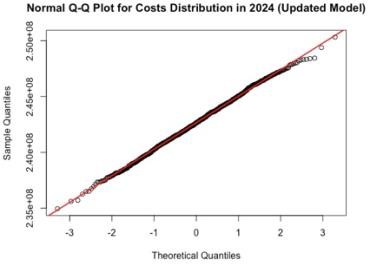
# Discussion

## Advantages and disadvantages of both models

An advantage of the updated model is that in the long run, the government will have a lower expected expenditure in regards to the benefit program. However, a disadvantage of this model is the expensive intervention costs of 20,000,000 Happyland dollars in the first two years, which, initially, makes the updated model more expensive than the current model.

An advantage in utilising the current model is that there are no intervention costs, and hence the government does not need to allocate any extra costs to the benefit program. This may allow the government to consider and use an expansionary fiscal policy, for example using tax cuts or increased government spending in infrastructure improvements which is also likely to benefit the citizens of Happyland. The primary disadvantages of implementing the current model are higher expected long run costs and a lower rate of transition from unemployed states to other states

The updated and current model’s Markov chain ergodic property generates an equilibrium situation where after a certain number of years, the number of citizens in each state will remain constant. This means that there will be fixed number of citizens that are unemployed (approximately 1111 and 1429 people for the updated and current model, respectively). Therefore, an advantage of implementing the updated model is that there will be fewer unemployed citizens in the long run when compared to the current model, which is optimal for the economy and the government.



***Figure*** 1.8

***Figure*** 1.7

Furthermore, the Shapiro-Wilk normality test was conducted using R studio, and the p-value for the cost series were greater than 0.05, and hence we cannot reject the null hypothesis of the data following a normal distribution.

# Expected long-run costs

The long run cost of both models further illuminates the difference in their effectiveness. As both models’ Markov chains are ergodic, the equilibrium situation was obtained by calculating the 100-step transition matrix for both Markov chains. The limiting probability property ensured that the distribution of citizens in each benefit program will remain constant and independent of the initial state, and the fixed distribution of citizens was multiplied with the vectorized payment plan, which results into the expected long run cost.

The expected long run cost for the current and updated model is: 205590062 and 194091787 Happyland dollars respectively. The annual long run cost for the current model is 5.56% more expensive than the updated model with intervention.

***Figure*** 1.9

# Discussion

## Advantages and disadvantages of both models

An advantage of the updated model is that in the long run, the government will have a lower expected expenditure in regards to the benefit program. However, a disadvantage of this model is the expensive intervention costs of 20,000,000 Happyland dollars in the first two years, which, initially, makes the updated model more expensive than the current model.

An advantage in utilising the current model is that there are no intervention costs, and hence the government does not need to allocate any extra costs to the benefit program. This may allow the government to consider and use an expansionary fiscal policy, for example using tax cuts or increased government spending in infrastructure improvements which is also likely to benefit the citizens of Happyland. The primary disadvantages of implementing the current model are higher expected long run costs and a lower rate of transition from unemployed states to other states

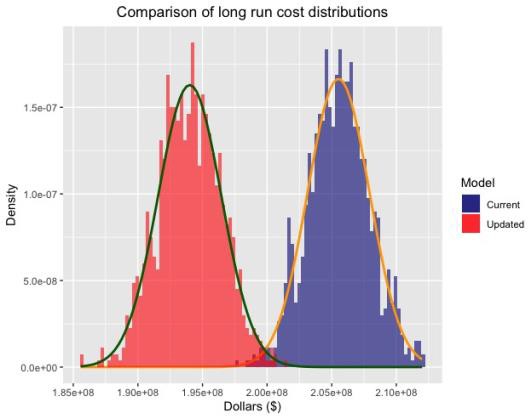
The updated and current model’s Markov chain ergodic property generates an equilibrium situation where after a certain number of years, the number of citizens in each state will remain constant. This means that there will be fixed number of citizens that are unemployed (approximately 1111 and 1429 people for the updated and current model, respectively). Therefore, an advantage of implementing the updated model is that there will be fewer unemployed citizens in the long run when compared to the current model, which is optimal for the economy and the government.

## Monte Carlo simulation and normal distribution

A statistical property shared by both models is that their costs distributions are normally distributed. This is advantageous to the government, as the tails of the cost distributions would be exponentially bounded (as a result of following a normal distribution), and will not have as many outliers with extremely high values (costs to the government), which is ideal in this business situation, as the likelihood that a significant deviation from the mean costs will be realised is minimised.

Furthermore, to investigate the density of long run costs generated from each model, it is assumed that the cost series follow a normal distribution. Monte Carlo simulation was conducted using R studio, where the integral of ∫𝜇+𝜎 𝑁(𝑥; 𝜇, 𝜎)𝑑𝑥 was calculated, with *x* being the random variable of long run costs. Approximately 68% of scores will lie within one standard deviation from the mean for both distributions, which is a favorable outcome for the government as their chances of paying ‘extreme’ amounts is reduced with a normal distribution.

𝜇−𝜎



***Figure*** 2.1

## Improvements

To further improve the current and/or updated model, the government may invest in a program that is more personalised to each citizen through conducting demographic analysis, particularly in collecting data on citizens’ socioeconomic information, such as education, residential area, marital status etc.

This information can be a factor in deciding the payment for each individual, and it allows for the benefit program to be more differentiated. Nevertheless, this model is likely to be very expensive for the government.

## Assumptions, limitations and constraints

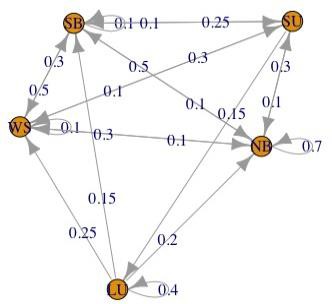
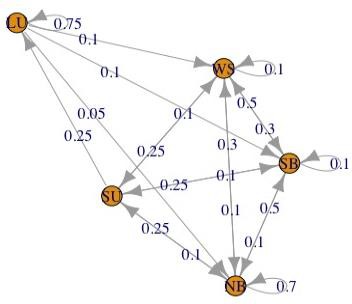
The analysis and explanation in the report has been made under several assumptions and are subjected to constraints:

* The data provided is limited to 10000 citizens, with the only information on citizens being their employment status.
* Inflation has not been taken into account when deciding the benefit payment as throughout the years, the payments is constant, but it will be worth less to the citizens as the years go on.
* It is assumed that uncertain events like a pandemic or war do not occur, as these major events may have a large-scale impact on the nation’s economy, and may result in contractionary/expansionary monetary and fiscal policy, which may change how the model is operated.
* It was assumed that the cost distribution for the current and updated model followed a normal (Gaussian) distribution.

## Conclusion

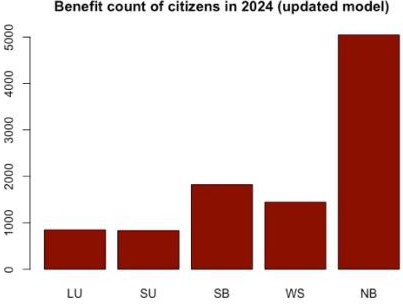
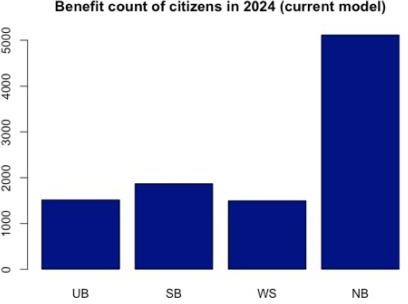
1000 simulations have been performed for both the current and updated model, with the given data of the initial state of benefit distribution. Both models demonstrate a fiscal policy that is successful in reducing the number of unemployed citizens for a given number of years, however, the updated model with intervention program outperforms the current model in relation to providing value to unemployed citizens, and reduced long run costs to the government. Therefore, it is recommended that the government invests in the updated model with the intervention program.

# Appendix



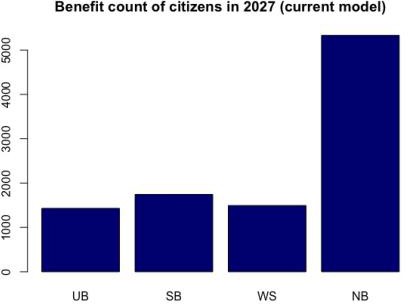
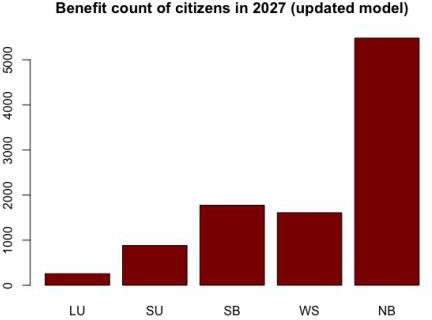
***Figure*** 2.2: State diagram of updated model without intervention

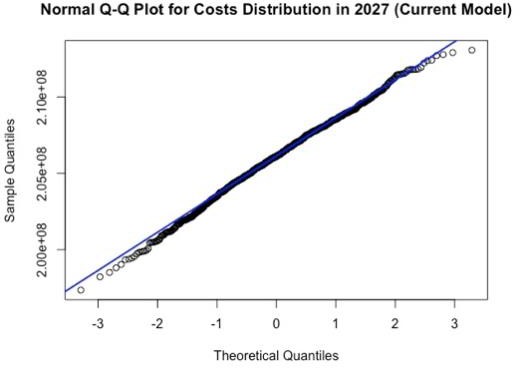
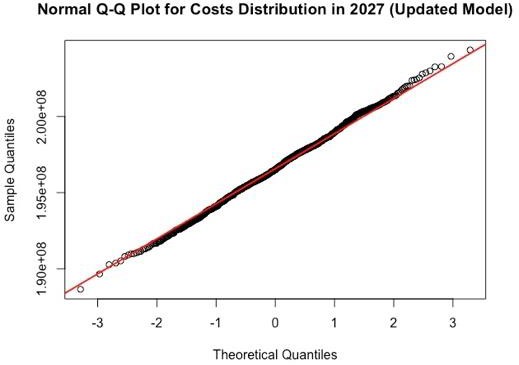
***Figure*** 2.3: State diagram of updated model with intervention



***Figure*** 2.4: Bar plot depicting the benefit count of citizens in 2024 (current model)

***Figure*** 2.5: Bar plot depicting the benefit count of citizens in 2024 (updated model)





***Figure*** 2.7: Bar plot depicting the benefit count of citizens in 2027 (updated model)

***Figure*** 2.6: Bar plot depicting the benefit count of citizens in 2027 (current model)

***Figure*** 2.9: Q-Q plot for updated model’s cost distribution in 2027

***Figure*** 2.8: Q-Q plot for current model’s cost distribution in 2027

# Bibliography

* + Ford, C, (2015), ‘Understanding Q-Q Plots’, | University of Virginia Library Research Data Services + Sciences, Data.library.virginia.edu, Available at:

<https://data.library.virginia.edu/understanding-q-q-plots/> [Accessed 15 July 2022].

* + Greenlaw, S. and Shapiro, D, (2022) ‘Using Fiscal Policy to Fight Recession,

Unemployment, and Inflation.’ , Opentextbc.ca, Available at:

<https://opentextbc.ca/principlesofeconomics2eopenstax/chapter/using-fiscal- policy-to-fight-recession-unemployment-and-inflation/> [Accessed 14 July 2022].

* + Hayes, A, Potters, C, and Beer, K, (2021), ‘What Is Demographics Analysis?’

Investopedia. Available at:

<https://[www.investopedia.com/terms/d/demographics.asp#:~:text=our%20edit](http://www.investopedia.com/terms/d/demographics.asp#%3A~%3Atext%3Dour%20edit) orial%20policies-

,What%20Are%20Demographics%3F,and%20death%20rates%2C%20and%20mor e.> [Accessed 14 July 2022].

* + Holtz, Y, (2022) ‘Basic R barplot customization’ , R-graph-gallery.com. Available at:

<https://r-graph-gallery.com/209-the-options-of-barplot.html> [Accessed 14 July 2022].

* + Kurt, W, (2015), ‘Monte Carlo Simulations in R — Count Bayesie.’ , Count Bayesie.

Available at: <https://[www.countbayesie.com/blog/2015/3/3/6-amazing-trick-](http://www.countbayesie.com/blog/2015/3/3/6-amazing-trick-) with-monte-carlo-simulations> [Accessed 14 July 2022].

* + Voskoglou, M., 2016. Applications of Finite Markov Chain Models to

Management<SUP>1</SUP>. [online] Article.sapub.org. Available at:

[<htt](http://article.sapub.org/10.5923.j.ajcam.20160601.02.html#Sec3)p[://article.sapub.org/10.5923.j.ajcam.20160601.02.html#Sec3>](http://article.sapub.org/10.5923.j.ajcam.20160601.02.html#Sec3) [Accessed 14

July 2022].

* + Covisum.com. (2022) ‘What is the difference between "heavy-tailed" and

Gaussian distribution models?’, Available at:

<https://[www.covisum.com/knowledge-base/what-are-heavy-tailed-](http://www.covisum.com/knowledge-base/what-are-heavy-tailed-) distributions> [Accessed 14 July 2022]