Extended Project Qualification Finley Cooper

What was the role of epidemic modelling in influencing government policies against COVID-19? A comparison of the policies used by England and Sweden in March to December 2020.

Contents

1 Abstract

The most powerful tool given to a government policy maker in the face of a pandemic is modelling. The modelling of an epidemic allows policy makers to predict the effectiveness of protective interventions without risk. This project focuses on non-pharmaceutical interventions (NPIs) in England and Sweden in the COVID-19 epidemic from March to December in 2020, broken down into these key sections:

- The production of epidemic models, including technical methods used and epidemic data to characterise the disease.
- The long-term strategies and subsequent policies chosen by England and Sweden to limit transmission.
- How did the models used by these two countries effect their strategy?
- Limitations of these models and a comparison with my own model.

The conclusion of my research is that epidemic modelling must be done throughout an epidemic, not just at the start, especially when stricter responses are enforced.

In my research I found a lack of precise data for important parameters for modelling COVID-19, especially data from early 2020. This led to inaccuracies in the data produced by modelling teams in both countries, which had a continued influence on policy implementation as there was little modelling of NPIs from the countries in the later stages of 2020. I saw how the policies implemented in England became much less thoughtful which hurt the UK's stricter strategy more than Sweden's strategy.

2 Introduction

Mathematical models of infectious diseases are a projection of how an infectious disease will progress over time, given a set of parameters. Policy makers can change these parameters to reproduce the effect of interventions which influences the spread of the disease, and the result from the model reflects which interventions might be the most successful.

Modelling diseases is a technique that has been used to influence public health policy since 1760, where Bernoulli analysed the influence of vaccination against smallpox (Bernoulli and Blower, 2004). The growth of compartmental models, pioneered by Ross in 1916, are models which is where the population is split into several classes. Ross' first models contained two classes, the unaffected, and the infected. Ross used differential equations to calculate the growth and decrease between each group over time, to graph the disease's progress (Ross and Hudson, 1917, p. 231).

Most of the first compartmental models were deterministic, which require massive populations. Realistically, this is not the case, so stochastic models are used instead, where randomness is introduced as people move between classes to calculate the probability of the progress of the epidemic (Bartlett, 1957).

Since then, models have been a crucial tool for policy makers, who now have access to computers, which allows models to become more complex and thus require more data to fit these new parameters (De Angelis et al., 2015). However even with all these parameters, many models solely rely on properties of the disease: they are not looking at how age groups interact differently, or public policy issues which can vary drastically country to country (Lewis and Al Mannai, 2021).

3 Making a Model

Making a model requires both data about the epidemic and a method of using this data to make predictions. To gain a better understanding of the methods and data collected by model makers in England and Sweden, I decided to research this and create my own model. Since the initial outbreak of COVID-19, more research has been published, so my model might be able to draw some conclusions missed by policy makers.

3.1 Data needed to model COVID-19

To model an infectious disease, data is needed that characterises the disease. This data must be accurate and reliable. Data collected in China, however, has been questioned by the international community (Campbell and Gunia, 2020), giving model creators little information to build a model with. High uncertainty levels also make stochastic modelling more appealing, as different probability distributions can be analysed within the range of values as was done by Ferguson et al. (2020a).

In *An Introduction to Infectious Disease Modelling*, Vynnycky and White (2010, p. 14) say that 'key features' of the disease must be identified to develop a model. The main characteristics they identify are the latent period (time between infection and onset of infectiousness), the infection duration, the basic reproduction number (R_0) which is a measure of how many people will be infected by one infected individual in a fully susceptible population, and how the infection affects age groups differently.

Most of these parameters can be measured through case tracing and studies of the infected. However, the percentage of asymptomatic infections is a key feature omitted by Vynnycky and White (Byrne et al., 2020), especially when considering policies, as asymptomatic individuals will not know if they are infected, and therefore will not abide by policies. When modelling, we can think of asymptomatic individuals as splitting the ‘Infected’ class into the asymptomatic individuals and symptomatic individuals, as asymptomatic individuals recover from COVID-19 quicker, so will move to the ‘Removed’ class at a faster rate (Byrne et al., 2020).

R_0 is harder to measure, as it must be estimated through modelling or calculation instead of pure measurement. This is shown in the range for R_0 in studies, which range from 0.4 to 12.58 (Dhungel et al., 2022). In early 2020, this was a larger problem, as the number of people infected with SARS-CoV-2 was so low that conclusive results were rare. Another reason for the large ranges in R_0 is shown in a meta-analysis of R_0, which found a disparity between the results of R_0 calculated using different methods (Liu et al., 2020), with one giving an average of 4.2 and another giving 2.44. This is an extremely significant difference. Using the fact epidemic will stop once the proportion of the population, *P*, has antibodies to prevent reinfection is given by,

$$P = 1 - \frac{1}{R_0}$$

With $R_0 = 4.2$, $P \approx 0.762 = \underline{76.2\%}$ and with $R_0 = 2.44$, $P \approx 0.590 = \underline{59.0\%}$

This difference between immunising half the population against immunising three quarters of the population is an immense difference in policy and reflects how important it is that data used in disease modelling must be accurate.

Table 1. Significant published estimates of data related to COVID-19.

Study	Date	R_0 results	Other notable results	Notes
Locatelli, Trächsel and Rousson (2021)	17/03/21	2.2 (95% CI: 1.9-2.6)		Gives the value for R_0 (2.2) as 'significantly lower' than R_0 in China
Li et al. (2020)	26/03/20	2.2 (95% CI, 1.4-3.9)	Mean incubation period of 5.2 days (95% CI: 4.1-7.0)	Used by Imperial College in Report 9.
Dhungel et al. (2022)	15/08/22	2.66 (95% CI, 2.41-2.94)		Does not agree that the Western Europe value for R_0 is 'significantly lower' than R_0 in China
Ma et al. (2021)	14/12/21	.	40.50% asymptomatic among the confirmed population with COVID	
Sayampanathan et al. (2021)	09/01/21		Symptomatic people had a 3.85 times higher incidence rate ratio	
Xin et al. (2021a)	12/06/21		Incubation period gamma distribution, Rate=0.61 P_{95}=13.1 Mean=6.3	

In Table 1, the results for R_0 vary widely study to study, with much wider confidence intervals for earlier studies, however all the studies reviewed agree that R_0 is somewhere between 2 and 3. However many of the studies use data from outside of Western Europe. This EPQ is focused on the value for R_0 in Western Europe, as it will be representative for England and Sweden which we assume have the same R_0.

The studies in the table, are conflicted if China and Western Europe have similar R_0 values. Locatelli, Trächsel and Rousson (2021) gives an estimated value R_0 in China of 3.32, compared to 2.2 for Western Europe. This is attributed to changes "on the social habits of a given population", but notes that published studies for R_0 in Europe for the time were rare, and difficulties existed with epidemiological data. Dhungel et al. (2022) gives a higher value for R_0 in the UK, 3.43, although the confidence intervals are exceptionally large (1.99-5.91).

As it cannot be confirmed that disease data can be transferred between countries, it is important that studies for calculating R_0 are done in the environment where the model will influence governmental policy, as even in 2022, Dhungel et al. only found 4 R_0 estimation studies for the UK suitable for meta-analysis, resulting in wide confidence intervals.

Poor data for R_0 transfers into uncertainty in the model outcomes, hurting policy implementations.

3.2 Model structure

Although some newer modelling methods have shown recent progress, compartmental models are the gold standard (Keeling and Eames, 2005). Much has changed since Ross' first compartmental models, where he noted that more classes must be considered to "represent the facts accurately" (Ross and Hudson, 1917, p. 239). However, we must choose what classes to use in the model and how the population can move between classes. This varies disease to disease, but for COVID-19, I will argue why the SEIR (Susceptible-Exposed-Infectious-Removed) model is the best.

The SIR model (Susceptible-Infectious-Recovered/Removed) is the base model for modelling infections, where once the infected individual has recovered, they are 'removed' from the system and cannot be reinfected (Vynnycky and White, 2010, p. 15). However, this model doesn't consider the period where the individual is 'exposed' to the infection, where they have the pathogen in their body, but the disease has not become prevalent enough to cause symptoms. This is the incubation period. The latent period is the period from infection to infectivity. The time between the latent period and the incubation period is important, as the individual will be unaware if they are infected, however are still transmitting the disease. A long incubation period will slow the spread of the disease.

Xin et al. (2021a) found a significant incubation period modelled with a Gamma distribution, mean 6.3 days and a rate of 0.61. As the pre-infectivity period is significant, our compartmental model should be extended to the SEIR model, where the host cannot transmit the disease in the exposed period.

Movement from the Removed class to the Susceptible class could also be considered, however paper from Stegger, et al. (2022) concluded that reinfection was 'rare' amongst the Omicron variant of SARS-CoV-2 so reinfection does not have to be considered, especially for the early variants of SARS-CoV-2 in 2020 (Vynnycky and White, 2010, p. 17).

3.3 Production of my model

For my own model, I used a modelling technique used by 3Blue1Brown (2020). The model uses collisions of small circles to represent contacts, with a chance of infection if an infectious person collides with a susceptible person. I chose this model because the visualisation makes debugging the program easier and lets me introduce policies easily. This makes the model an Individual-based model, as we are tracking the state of each individual and each infection is tracked and determined separately (Vynnycky and White, 2010, p. 150). This gives more control over the model, as we can change the infection chance in various locations.

Before introducing a disease, I had to create the population and town to infect. The town was created using UK census data, which may cause bias towards England over Sweden, however the difference shouldn't be massive, as they are both similar Western European countries. I fit the household composition (Appendix 1) to the ages of the population (Appendix 2) to create the population and housing. I gave out jobs using government job data (Appendix 3) and gave out schedules for people to go to events, such as the theatre, shops, and other recreational activities, which would be what policies will mainly act on.

To introduce COVID-19, I tagged each person with their class in the SEIR model, then infected three individuals at the simulation's start. To fit the model's R_0 to COVID-19, I set up the simulation with everyone infected, and I counted the number of 'would be' successful infections until the entire population had recovered. Then I averaged the number of successful infections per person to calculate R_0. I ran the simulation with different transmission chances, until I found a value that gave a R_0 of 2.56 (95% CI: 2.56-2.69) which was suitable for the simulation.

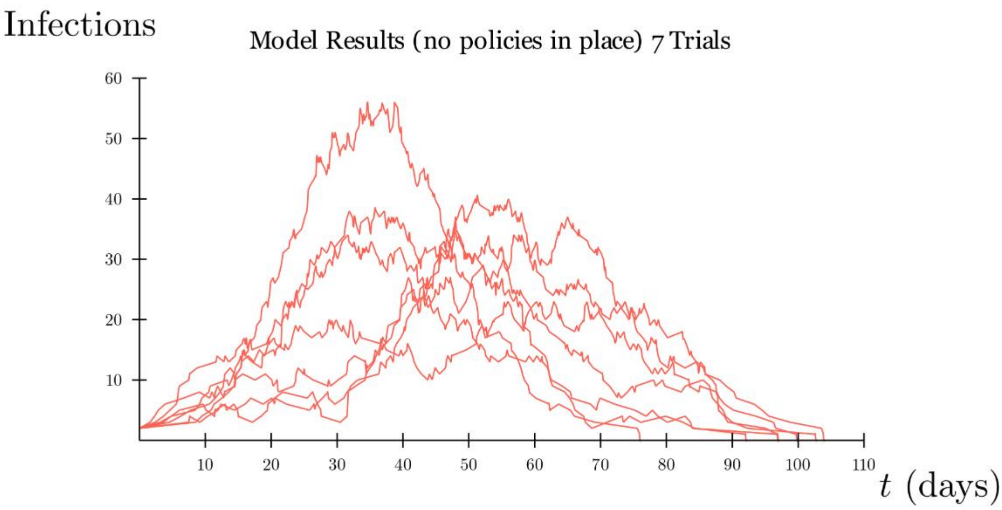
To calculate the incubation period for each person, I sampled from the gamma distribution described in Table 1 (Xin et al. 2021a). The mean time between the latent period and the incubation period is 1.4 days (Xin et al. 2021b), so I will take the latent period as a constant 1.4 days less than the sampled incubation period as I cannot take a sample for the latent period because the durations are not independent.

I struggled to find data on the infectious period, for both symptomatic and asymptomatic people. A literature review of the infectious period by Byrne et al. (2020) found a mean of 13.4 days for the post-symptom onset infectious period and a range of 6.5-9.5 days for asymptomatic individuals. However, the data is limited, as the author notes "substantial variation" in the estimates. Furthermore, my model assumes infectiousness remains constant as the infected person recovers, so I chose a post-symptom onset infectious period of 10 days, which gives 11.4 days for the entire infectious period. The figure shared by Byrne et al. (2020) for the duration of the infectious period for asymptomatic infections does not agree on a value at all. I will use 6 days as an estimate.

Many studies disagreed on the percentage of the population infected who were asymptomatic, as very early figures only gave a single child being asymptomatic in the first 425 cases of COVID-19 in Wuhan (Li et al. 2020), but a meta-analysis on 95 studies gave a pooled percentage of asymptomatic infectious cases of 40.50% (Ma et al. 2021).

Finally, I decreased the chance of infection from asymptomatic individuals, as a seroprevalence study found that symptomatic individuals infect at a 3.85 times larger rate (Sayampanathan et al., 2021).

Figure 1. Control runs of the model with a population of 162.



The results show a clear epidemic 'peak' in all trials with substantial variation between each trial. The general epidemic curve is visible in each trial and the total proportion of the population infected is not outrageously low or high, so I deemed the model fit to implement policies on.

4 Exploring Policies Introduced

With a better understanding of the creation of an effective model, I chose to investigate further the policies used by England and Sweden, as they had different responses to COVID-19. England's policy followed a 'suppression' strategy, where policies are introduced to keep the reproduction number, below 1, reversing the growth of the epidemic. This involves implementing lockdowns and quarantining of households. Sweden instead followed the 'mitigation' technique, which includes advisory policies for social distancing, and contact tracing (Ferguson et al. 2020a) .

Yan et al. (2020) investigates the differences between decentralised regimes and centralised regimes in the difference in policy between Sweden and France. The Swedish response follows a 'Nudge' strategy with suggestive and weak policies, and England's response fits the centrally enforced 'Decree' strategy, like France and most other Western European countries. These strategies lead to different epidemic curves and follow similar ideas to the suppression and mitigation strategies in Ferguson et al. (2020a).

It should be noted that the success of policies is very hard to measure, as other factors such as population density, age distribution, and public compliance with policies all contribute to the number of infections and deaths in an epidemic, so success cannot be determined with certainty. I am more focused in understanding how modelling is used best by each country, not the overall success, as that is far beyond the scope of this project.

4.1 Policies Introduced by Sweden

At the beginning of the pandemic there was a lot of controversy surrounding the Swedish response, with The Daily Mail saying that Sweden was heading for a 'catastrophe' (Connolly, 2020). The country was accused of trying to achieve 'herd immunity', which would have caused countless deaths, however I disagree with these initial opinions, as the Swedish response gave the population more protection against multiple future waves.

Many observers were quick to disagree with the Swedish response to COVID-19, due to it having one of the lowest COVID-19 response stringency indexes (Hale et al. 2020), however the Swedish Foreign Minister, declared that the response to the pandemic was a “marathon not a sprint”, therefore a less strict policy was chosen to make sure it was acceptable for a long period (Heath, 2020).

I’ve discussed that proportion of the population who need to be immune against SARS-CoV-2 to prevent epidemic growth is around 60%, depending on the value of R_0 used. Adding the policies implemented by Sweden, which include a ban on unessential travel and social distancing advisories (Yan et al. 2020), an infection of 60% of the population doesn’t seem like the strategy that Sweden was going for, which is what would be needed to achieve herd immunity.

WHO Director-General Ghebreyesus (2020) remarked that,

> “Never in the history of public health has herd immunity been used as a strategy for responding to an outbreak, let alone a pandemic.”

Instead, the Swedish policy was a policy to not give herd immunity, but herd protection, preventing multiple waves. This can be shown as Sweden never enforced a ban on education for those under 16 (Halin et al. 2020), using the younger generations, who had a much lower fatality rate, to safely increase the proportion of the population who are immune. This is also supported by a skew in infections towards yonger people (Monod et al. 2021). This high level of immunity against COVID-19 in more mobile populations, could be an explanation of the unexpectedly low hit rate of Omicron in Stockholm County, which fell below modelled predictions of a 60% hit rate to 30% (Carlsson and Söderberg-Nauclér, 2022), suggesting that this was a result of the herd protection policy.

4.2 Policies Introduced by England

The policy introduced by the UK was a typical ‘suppression’ strategy; the goal was to keep the rate of infections in the population decreasing for as long as the virus is circulating in the population until a vaccine becomes available (Ferguson et al. 2020a). However, the UK attempted to implement ‘adaptive policy’ where policies are only introduced after some metric of the epidemic reaches a threshold, which, although very effective in theory, did not result in the UK being protected against fatalities and nationwide disruption due to COVID-19, due to the confusing implementation of policies.

The Imperial College COVID-19 Response Team published the report Ferguson et al. (2020a), which investigated the impact of intervention strategies to reduce COVID-19 mortality, and the results directly informed policymaking in the UK. The report concludes with social distancing being the most significant policy, followed by home isolation of confirmed cases. The report also presented an ‘adaptive policy’ where some policies, such as social distancing and school closures, are only enforced when hospitals were under stress. It was also mentioned in the discussion section that local policies were more efficient than national policies.

The first national lockdown was introduced from March to April 2020 where all but essential travel was prohibited, enforced with the closure of non-essential facilities (Press Association Reporters, 2020).

This first lockdown was simple and understandable, however from April to September, the counties of the UK began to take different approaches to exit the lockdown. Scotland announced a zero-COVID policy (Sridhar and Chen, 2020), which contrasted the “modest” lifting of measures in Wales, and the static situation in England (Roderick, 2020). This was made more confusing with the creation of a three tier COVID-19 alert level, which replaced the piecemeal implementation of more stringent measures, the most notable being the lockdown of Leicester. I believe that the COVID-19 alert levels were the government’s attempts to implement the local policies praised in Ferguson et al. (2020a). However, these techniques were shortly replaced with another national lockdown throughout November (Kuenssberg and Ghosh, 2020), further adding to the confusion.

Overall, the English policy followed the ‘Decree’ strategy, but devolved into confusion throughout 2020 as the transmission of SARS-CoV-2 progressed through the population, and forced the government to be overreactive to changes, reflected in a complex policy structure.

5 Usage of Models

Both England and Sweden used models to provide information to policy makers, however both countries took varied approaches to modelling, with both countries using similar models, but commissioning this research in different ways, with the British data being commissioned for the government in one large model for the entire of the UK, and even some other countries, compared to the decentralised Swedish models which come in many shapes and forms, and concentrate on the analysis of more specific issues in a specific region of Sweden.

5.1 Usage of Models by Sweden

The models were found on a Swedish government page, where there is a reference to a GitHub profile with the models developed by the Public Health Agency of Sweden (Folkhälsomyndigheten, nd).

From looking at the models, it's clear that the models focus on 'snapshot' predictions of the disease, given the current progress of the disease, without implementing any NPIs in the model. Namely, an estimation of the number of infected individuals in four regions of Sweden, led by Brouwers (2020), uses an SEIR model, however the model is deterministic instead of stochastic, which is less effective for policy making, not letting the Swedish government fully assess the risk of the epidemic. The model also calculates the future progress of the disease from existing incident cases. The data used was collected over a period of four months, however the epidemic was still in its initial stages, with Stockholm being the only region selected with more than one hundred cases a day. The model also predicts a large secondary epidemic wave if restrictions were lifted.

The conclusions of the model are hard to figure out. There is no mention in the report of if the predictions are acceptable losses to bear, if the government should change their response due to the results, or even if the report was seen by policy makers. This was an issue throughout the pandemic, with a similar point raised in the long-awaited government-initiated commission called the Coronakommissionen (Bouder et al. 2022), which was tasked in evaluating the government actions implemented to limit the spread of SARS-CoV-2. The final report of the Coronakommissionen was critical of the Public Health Agency, mentioning that the data given to the government was not enough to support a change in policy in November 2020. It is also mentioned that the communications from the Public Health Agency to the public were much too general and left much to interpretation. Although not discussed in the report, I think that this is because the Public Health Agency failed to investigate sufficiently into the effects of NPIs, and therefore couldn't inform the public with precise instructions.

One other interesting point raised by the Commission was surrounding the snapshots of the epidemic's progression developed by the Public Health Agency. The Commission says the government does not have the authority or the ability to judge on scientific controversies, which in this case is the type of policy implemented by Sweden, as it is task that has been designated to its expert agencies, like the Public Health Agency. This is an important fact, as the risk assessments sent to the government by the Agency were seen without full context of the extremely incomplete and uncertain data we've just discussed. It's possible that this caused the government to double-down on the 'Nudge' strategy, which could have been disastrous due to the high levels of uncertainty.

5.2 Usage of Models by England

Compared to the usage of models used by Sweden, the UK initially took a more rigorous approach, however, throughout the progression of the epidemic in late 2020, the government starts failing to implement recommendations from the Scientific Advisory Group of Emergencies (SAGE), which lead to a surge of cases, breaking apart the suppression strategy, however this was much more disastrous for the UK over Sweden.

One of the most prominent models for COVID-19, was CovidSim, a stochastic individual-based model developed by the Imperial College COVID-19 Response Team led by Ferguson (2020b). This model was significant as it gave a prediction of 510,000 deaths due to SARS-CoV-2 in the UK and caused the UK government to backtrack from a policy which focused on building 'herd immunity' to a suppression strategy (Bostock, 2020).

As expected, much of the data used to model COVID-19 in Ferguson's first model was incorrect, most significantly being a too high infectivity of asymptomatic individuals at 67% of symptomatic individuals, whereas the source shown in Table 1, puts the percentage at around 26% of symptomatic individuals. As asymptomatic infections accounted for a third of cases in the release of CovidSim, we could have seen a pessimistic result in Report 9. This data could have been adjusted throughout the pandemic, bringing new advice, especially with variants of SARS-CoV-2, to allow the safe lifting of some restrictions, where policy was seen as strict enough.

The government was very vocal about following 'the science' in policy making, however in late September, the government refused to act on a proposal from the Scientific Advisory Group of Emergencies (SAGE) that recommended an immediate short and tough lockdown based on new forecasts from a different model (Tighe, et al 2020). This resulted in a second epidemic wave and forced the government to shortly reintroduce many of the measures which had been slowly lifted in the previous months.

This was the first time that the government had directly gone against Ferguson's initial research, which clearly set out that measures were only to be lifted once immunity had reached a high enough level through vaccination (Ferguson et al. 2020a). Lifting policies before would only move the country backwards, as the UK had a much lower antibody presence than Sweden which introduced minimal protections. Therefore, Sweden could afford to ignore the Public Health Agency, as the next epidemic wave would be much lighter, as a significant proportion of the population were immune to SARS-CoV-2, as seen in the low hit rate later in Stockholm (Carlsson and Söderberg-Nauclér, 2022). This would not work for the UK however, as the suppression strategy relies on a vaccine being produced to provide immunity.

Although SAGE did have a subgroup for modelling, the SPI-M-O, their models were limited. They didn't discuss the influence of any new

NPIs, they mostly published 'short-term-forecasts' and the reports published didn't link together to provide clear recommendations. SAGE has also been criticised for creating an environment for 'scientific groupthink' (Coker, 2020), due to too much familiarity between members. This explains SAGE's unremarkable conclusions, who might have not provided the atmosphere for prominent reports.

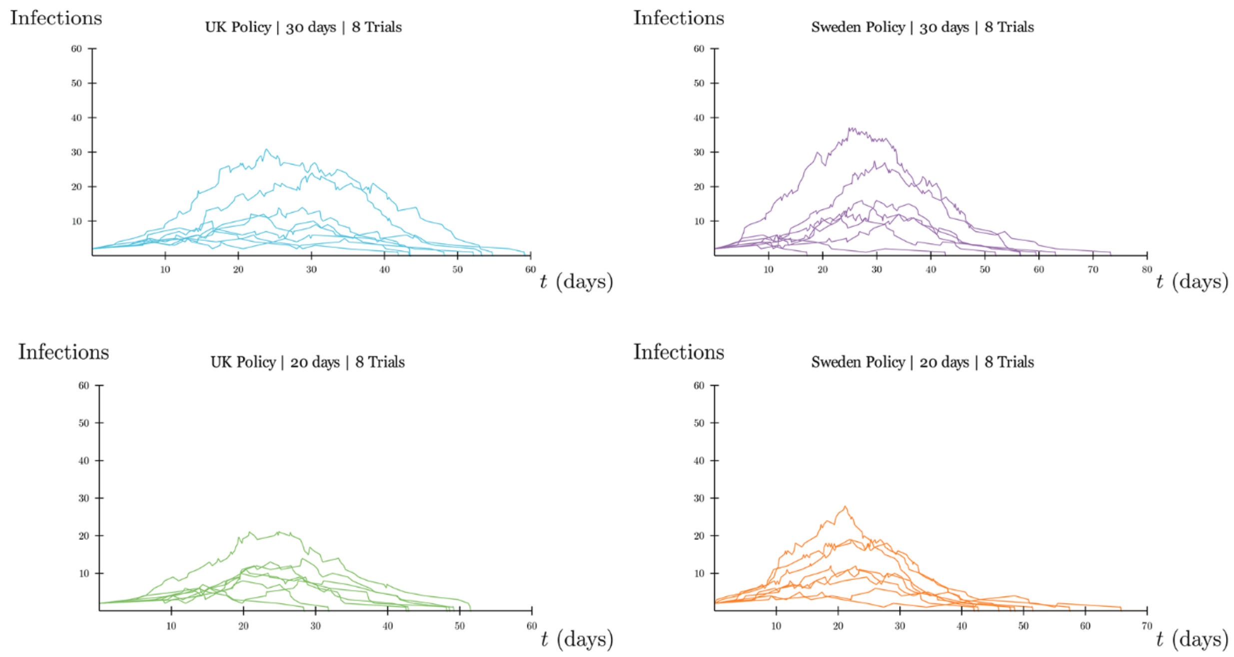
6 Analysis Against my own Model

Throughout this project, I’ve discussed how neither the UK nor Sweden fully utilised the power of disease modelling in the COVID-19 pandemic. In this section I will analyse the models discussed in this essay, using data created by my model.

The main factor omitted in all the models I’ve looked at (and my own model to an extent) is time. For example, Ferguson’s Report 9 only models up to 2021 in Figure 4, where the adaptive policy is shown. This is a justifiable choice, as much can change in year, such as a new strain, new policy choices, or seasonal changes; however, the government didn’t follow up another report from Imperial College about the effects of NPIs in the UK. The short-term forecasts by the SPI-M-O were not enough to ensure a change in policy would have a positive impact on the country, leading to the government removing policies without any scientific evidence. This is like Sweden, where the Public Health Agency also omitted modelling research into the effects of NPIs. This wasn’t as large of a problem in Sweden however, as the ‘herd protection’ policy’s success was independent from the length of the epidemic, however in the UK, the ‘Decree’ strategy was very dependent on national vaccination to allow policy to be removed.

Another element of time that should have been investigated before the pandemic should have been how quickly policy should be implemented, in my model, only a 10-day additional delay between the first case and policy implementation caused a massive decrease in infections, shown in Figure 2. With hindsight, it’s easy to question why countries took so long to implement strict policies against COVID-19, however at the time it wasn’t clear how the pandemic was going to progress. This further proves the importance of modelling to predict how dangerous a virus will be before it has fledged into an epidemic. I still think that Ferguson should have investigated the effects of implementing policy quicker, to get the government to act quicker if a second wave was to occur.

Figure 2. Results of my model. The number of days is the time between the first infection and policy implementation.



There are several drawbacks to my model, the most significant problem is that the population size is too small, at 162, it's easier for the disease the die out in the population than in an entire country. I was only able to run 8 trials per scenario which makes the data hard to reproduce or make confident conclusions from. Both problems are caused by my model being too computationally heavy, due to the Manim engine, which is not designed for long animations which my model uses for calculations. Even with all of this, it is still clear that implementing a policy quickly is much more important than the small additions shown by the bottom two graphs having a much lower maximum number of infections and a lower total number of infections. This is because the social distancing policy and self-isolating policy are the most effective policies in my simulation, shown also in Ferguson's model (Ferguson et al. 2020a).

For more details on my model, see Appendix 4

7 Conclusion

The role that epidemic modelling has in the development of responses to COVID-19 is an important metric to understand. It has been shown that it is vital to utilise models, creating a clear channel of communication between epidemiologists and government policy makers, without introducing an overreliance on certain produced data, especially when that data is flawed due to unseen factors in the introduction of policies, or when new epidemic data comes to light.

In my analysis, I think that both England and Sweden did not fully use epidemic modelling throughout the epidemic. The models produced by England, most significantly Ferguson's (2020a), produced clear results. However, the UK government didn't produce any other similarly significant material and sometimes disregarded warnings from SAGE's modelling team. The Swedish models did not include long term forecasts, and unlike in England the conclusions in the reports were unclear and did not give clear advice to the Swedish government, leading the government to be more tentative when enacting restrictions.

Although clearly Sweden was hit less by the pandemic than England, this doesn't say much about how the modelling was used by both countries. The Swedish policy was based of little information, which could have been disastrous for Sweden if the disease spread faster than estimates showed. The British policy was based on an exact modelling report and was successful at the start, however both countries decreased research into the effect of NPIs in the pandemic, which led to rushed and possibly dangerous policy implementation near the end of 2020. This had a worse effect on England rather than Sweden however, as the herd protection that Sweden had built in their younger population decreased the rate of infection. The importance of speed when implementing models was also not stressed enough, which was picked up by the Coronakommission and myself.

Overall, I would say that England made a more effective use of epidemic modelling than Sweden, although both countries did not use it enough after the start of the epidemic.

Bibliography

Resources used in model creation

Vynnycky, E., White, R.G. and Fine, P.E.M. (2011) An introduction to infectious disease modelling. Oxford: Oxford University Press.

Locatelli, I., Trächsel, B. and Rousson, V. (2021) “Estimating the basic reproduction number for COVID-19 in Western Europe,” PLOS ONE, 16(3). Available at: https://doi.org/10.1371/journal.pone.0248731.

Byrne, A.W. et al. (2020) Inferred duration of infectious period of SARS-COV-2: Rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases, BMJ Open. British Medical Journal Publishing Group. Available at: https://bmjopen.bmj.com/content/10/8/e039856 (Accessed: April 10, 2023).

Dhungel, B. et al. (2022) “Reliability of early estimates of the basic reproduction number of COVID-19: A systematic review and meta-analysis,” International Journal of Environmental Research and Public Health, 19(18), p. 11613. Available at: https://doi.org/10.3390/ijerph191811613.

Liu, Y. et al. (2020) “The reproductive number of COVID-19 is higher compared to SARS coronavirus,” Journal of Travel Medicine, 27(2). Available at: https://doi.org/10.1093/jtm/taaa021.

Li, Q. et al. (2020) “Early Transmission Dynamics in Wuhan, China, of novel coronavirus–infected pneumonia,” New England Journal of Medicine, 382(13), pp. 1199–1207. Available at: https://doi.org/10.1056/nejmoa2001316.

Ma, Q. et al. (2021) “Global percentage of asymptomatic SARS-COV-2 infections among the tested population and individuals with confirmed covid-19 diagnosis,” JAMA Network Open, 4(12). Available at: https://doi.org/10.1001/jamanetworkopen.2021.37257.

Sayampanathan, A.A. et al. (2021) “Infectivity of asymptomatic versus symptomatic COVID-19,” The Lancet, 397(10269), pp. 93–94 (Supplementary Material). Available at: https://doi.org/10.1016/s0140-6736(20)32651-9.

Xin, H. et al. (2021a) “The incubation period distribution of Coronavirus Disease 2019: A systematic review and meta-analysis,” *Clinical Infectious Diseases*, 73(12), pp. 2344–2352. Available at: https://doi.org/10.1093/cid/ciab501.

Xin, H. et al. (2021b) “Estimating the latent period of coronavirus disease 2019 (COVID-19),” *Clinical Infectious Diseases*, 74(9), pp. 1678–1681. Available at: https://doi.org/10.1093/cid/ciab746.

3Blue1Brown (2020) Simulating an epidemic [Online video]. Available at: https://www.youtube.com/watch?v=gxAaO2rsdIs (Accessed: April 16, 2023).

Individual census data for creating the model in Appendix 1, 2, and 3

Resources used in policy investigation

Campbell, C. and Gunia, A. (2020) Can we believe any of China's coronavirus numbers? Time. Available at: https://time.com/5813628/china-coronavirus-statistics-wuhan/ (Accessed: April 10, 2023).

Bernoulli, D. and Blower, S. (2004) “An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it,” Reviews in Medical Virology, 14(5), pp. 275–288.

Available at: https://doi.org/10.1002/rmv.443. Translated and republished by Sally Blower from the original paper published in 1766 by Bernoulli.

Dietz, K., Heesterbeek, J. Bernoulli was ahead of modern epidemiology. Nature 408, 513–514 (2000). https://doi.org/10.1038/35046270

Ross, R. and Hudson, H. (1917) “An application of the theory of probabilities to the study of a priori pathometry. —Part III,” Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 93(650), pp. 225–240. Available at: https://doi.org/10.1098/rspa.1917.0015.

De Angelis, D. et al. (2015) “Four key challenges in infectious disease modelling using data from multiple sources,” Epidemics, 10, pp. 83–87. Available at: https://doi.org/10.1016/j.epidem.2014.09.004.

Yan, B. et al. (2020) “Why do countries respond differently to COVID-19? A comparative study of Sweden, China, France, and Japan,” The American Review of Public Administration, 50(6-7), pp. 762–769. Available at: https://doi.org/10.1177/0275074020942445.

Lewis, T.G. and Al Mannai, W.I. (2021) “Editorial: Modeling epidemics - why are models wrong?” Frontiers in Public Health, 9. Available at: https://doi.org/10.3389/fpubh.2021.754746.

Bartlett, M.S. (1957) “Measles periodicity and community size,” Journal of the Royal Statistical Society. Series A (General), 120(1), p. 48. Available at: https://doi.org/10.2307/2342553.

Tembine, H. (2020) “Covid-19: Data-driven mean-field-type game perspective,” Games, 11(4), p. 51. Available at: https://doi.org/10.3390/g11040051.

Ferguson, N. et al. (2020a) Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. Available at: https://doi.org/10.25561/77482 (Accessed: April 5, 2023).

Ferguson, N. et al. (2020b) COVID-19 CovidSim Model [Source code]. Available at: https://github.com/mrc-ide/covid-sim (Accessed: April 13, 2023).

Keeling, M.J. and Eames, K.T.D. (2005) “Networks and epidemic models,” Journal of The Royal Society Interface, 2(4), pp. 295–307. Available at: https://doi.org/10.1098/rsif.2005.0051.

Stegger, M. et al. (2022) Occurrence and significance of Omicron Ba.1 infection followed by BA.2 reinfection, medRxiv. Cold Spring Harbor Laboratory Press. Available at: https://www.medrxiv.org/content/10.1101/2022.02.19.22271112v1 (Accessed: April 13, 2023).

Byrne, AW. et al. (2020) .Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. Available at: https://doi.org/10.1136/bmjopen-2020-039856

Hale, T, et al. (2020). ““Variation in Government Responses to COVID-19” Version 14.1. Blavatnik School of Government Working Paper. 27 July 2022. Available at: www.bsg.ox.ac.uk/covidtracker

Heath, R. (2020) Swedish minister defends coronavirus approach, shrugs off far-right embrace, POLITICO. POLITICO. Available at: https://www.politico.eu/article/sweden-coronavirus-leader-ann-linde-defends-approach-shrugs-off-far-right-embrace/ (Accessed: April 21, 2023).

Connolly, P., (2020). Is softly, softly Sweden heading for catastrophe. The Daily Mail. Available at: https://www.dailymail.co.uk/debate/article-8173691/Is-softly-softly-Sweden-heading-catastrophe.html (Accessed: April 21, 2023)

Ghebreyesus , T.A. (2020) Who director-general's opening remarks at the media briefing on COVID-19 - 12 October 2020, World Health Organization. World Health Organization. Available at: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---12-october-2020 (Accessed: April 21, 2023).

Carlsson, M. and Söderberg-Nauclér, C. (2022) “Covid-19 Modeling outcome versus reality in Sweden,” Viruses, 14(8), p. 1840. Available at: https://doi.org/10.3390/v14081840.

Bostock, B. (2020) How 'professor lockdown' helped save tens of thousands of lives worldwide - and carried covid-19 into Downing Street, Business Insider. Business Insider. Available at: https://www.businessinsider.com/neil-ferguson-transformed-uk-covid-response-oxford-challenge-imperial-model-2020-4?r=US&IR=T (Accessed: May 1, 2023).

Hallin, A.E. et al. (2020) “No learning loss in Sweden during the pandemic: Evidence from primary school reading assessments,” International Journal of Educational Research, 114, p. 102011. Available at: https://doi.org/10.1016/j.ijer.2022.102011.

Monod, M. et al. (2021) “Age groups that sustain resurging COVID-19 epidemics in the United States,” Science, 371(6536). Available at: https://doi.org/10.1126/science.abe8372.

Press Association Reporters (2020) Timeline of key coronavirus events since the UK was put into Lockdown Six Months Ago, The Independent. Independent Digital News and Media. Available at: https://www.independent.co.uk/news/uk/home-news/coronavirus-uk-timeline-lockdown-boris-johnson-pubs-test-and-trace-vaccine-b547630.html (Accessed: April 28, 2023).

Sridhar, D. and Chen, A. (2020) “Why Scotland’s slow and steady approach to covid-19 is working,” BMJ, p. m2669. Available at: https://doi.org/10.1136/bmj.m2669.

Roderick, V. (2020) Coronavirus: 'modest' lockdown changes announced in Wales, BBC News. BBC. Available at: https://www.bbc.co.uk/news/uk-wales-52584690 (Accessed: April 28, 2023).

Kuenssberg, L. and Ghosh, P. (eds) (2020) Covid-19: PM announces four-week England lockdown, BBC News. BBC. Available at: https://www.bbc.co.uk/news/uk-54763956 (Accessed: April 28, 2023).

Tighe, C. et al (2020) “Johnson gambles by splitting from his scientists,” Financial Times. Available at: https://www.ft.com/content/af637ae6-0219-49f1-bce8-8bef4ef7000a (Accessed: May 2, 2023)

Folkhälsomyndigheten (no date) Analyser och prognoser om covid-19 [website]. Available at: https://www.folkhalsomyndigheten.se/smittskydd-beredskap/utbrott/aktuella-utbrott/covid-19/statistik-och-analyser/analys-och-prognoser/ (Accessed: 03 May 2023).

Brouwers, L (2020) “Estimates of the number of infected individuals during the covid-19 outbreak in the Dalarna region, Skåne region Stockholm region, and Västra Götaland region, Sweden” Folkhälsomyndigheten. Available at: https://www.folkhalsomyndigheten.se/publikationer-och-material/publikationsarkiv/e/estimates-of-the-number-of-infected-individuals-during-the-covid-19-outbreak/ (Accessed: May 3, 2023)

Bouder, F. et al. (2022) “Coronakommission Final report” Coronakomission. Available at: https://coronakommissionen.com/publikationer/slutbetankande-sou-2022-10/ (Accessed: May 3, 2023)

Coker, R. (2020) Coronavirus can only be beaten if groups such as sage are transparent and accountable, The Guardian. Guardian News and Media. Available at: https://www.theguardian.com/commentisfree/2020/apr/27/coronavirus-sage-scientific-groupthink (Accessed: May 6, 2023).

Appendices

Appendix 1 – Household census data analysis [Python 3.10.7]

```
# File: data_analysis/census/households.py
# Household characteristics from the 2021 Census
# Source of data:
https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmedi
ausage/bulletins/householdandresidentcharacteristicsenglandandwales/census2021
# Office for National Statistics - Census 2021
# All data is rounded to the nearest 0.1%

DATA = {
    "one_person_aged_66_or_over": 12.9,
    "one_person_other": 17.3,
    "couple_family_no_children": 16.7,
    "couple_family_dependent_children": 18.8,
    "couple_family_all_nondependent_children": 6.4,
    "lone_parent_dependent_children": 6.9,
    "lone_parent_all_nondependent_children": 4.2,
    "other": 16.8,
}

# We will not consider the "other" category, so we will remove it from the data and scale up the other
categories so the total is 100%
# This could lead to a possible error, as the other category is heavily biased towards one-person
households (students).
# so our data will overrepresent families, however this is not a problem for our purposes as we are
not simulating a university or college campus.

for key in DATA:
    DATA[key] *= 100 / (100 - DATA["other"])

del DATA["other"]

# Family households 2021 Dataset
#
https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/datasets/familie
sandhouseholdsfamiliesandhouseholds
# Office for National Statistics - Census 2021

# This data set is used to determine the number of children in a household
# The data is split into 3 categories: 0 children, 1 child, 2+ children
# As 0 children is already accounted for in the household data, we will only use the other two
categories
# We will assume that single parent households have the same number of children as couple households
as this is not covered in the data set
# Out of the 24510 family households with dependent children
# The number of households with 1-2 children is 19133
# The number of households with 3+ children is 5377

# We will assume that no households have more than 3 children
# We will also assume that the number of households with 1 child is the same as the number of
households with 2 children
# This is not true, but it is a reasonable assumption for our purposes

# Breaking up the couple family households with dependent children into 1 child and 2+ children

DATA["couple_family_one_child"] = DATA["couple_family_dependent_children"] * (19133 / 2) / (19133 +
5377)
DATA["couple_famliy_two_children"] = DATA["couple_family_dependent_children"] * (19133 / 2) / (19133 +
5377)
DATA["couple_family_three_children"] = DATA["couple_family_dependent_children"] * 5377 / (19133 +
5377)

# Then the same for the lone parent

DATA["lone_parent_one_child"] = DATA["lone_parent_dependent_children"] * (19133 / 2) / (19133 + 5377)
DATA["lone_parent_two_children"] = DATA["lone_parent_dependent_children"] * (19133 / 2) / (19133 +
5377)
DATA["lone_parent_three_children"] = DATA["lone_parent_dependent_children"] * 5377 / (19133 + 5377)
```

```
# We will then assume that households with only non-dependent children only have 1 child
DATA["couple_family_one_nondependent_child"] = DATA["couple_family_all_nondependent_children"]
DATA["lone_parent_one_nondependent_child"] = DATA["lone_parent_all_nondependent_children"]


# Finally one_person_other will be simplified to one_person
DATA["one_person"] = DATA["one_person_other"]

# We will then remove the old categories
del (
    DATA["couple_family_dependent_children"],
    DATA["couple_family_all_nondependent_children"],
    DATA["lone_parent_dependent_children"],
    DATA["lone_parent_all_nondependent_children"],
    DATA["one_person_other"],
)

print(DATA)
# File: project/Population/households.py
# Results from data_analysis/census/households.py (Appendix 1)
DATA = {
    "one_person_aged_66_or_over": 15.504807692307692,
    "couple_family_no_children": 20.072115384615383,
    "couple_family_one_child": 8.819506559332142,
    "couple_famliy_two_children": 8.819506559332142,
    "couple_family_three_children": 4.9571407274895645,
    "lone_parent_one_child": 3.2369465563506266,
    "lone_parent_two_children": 3.2369465563506266,
    "lone_parent_three_children": 1.8193761180679786,
    "couple_family_one_nondependent_child": 7.6923076923076925,
    "lone_parent_one_nondependent_child": 5.048076923076923,
    "one_person": 20.79326923076923,
}


def distribute_households(number_of_households):
    # Distribute the 70 houses across the 7 categories by the value given in DATA
    housesholds = {}

    for key in DATA:
        housesholds[key] = int(DATA[key] / 100 * number_of_households)

    # Add the remainder to the category with the highest percentage
    housesholds[max(housesholds, key=housesholds.get)] += number_of_households -
sum(housesholds.values())

    # check if the sum of the values is equal to NUMBER_OF_HOUSEHOLDS
    assert sum(housesholds.values()) == number_of_households

    return housesholds
```

Appendix 2 – Age census data analysis [Python 3.10.7]

```
# File: project/Population/ages.py
# Percentages of the popluation of England and Wales for each age.
# Source of data:
https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/dataset
s/ageingpopulationestimates
# Office for National Statistics - Census 2021
# First age is aged under 0, last is those aged 100 and over
# All data is rounded to the nearest 0.1%
# The total population of England and Wales recorded in the 2021 Census was 59,597,542

DATA = [1.0, 1.1, 1.1, 1.1, 1.1, 1.2, 1.1, 1.2, 1.2, 1.2, 1.2, 1.2, 1.2, 1.2, 1.2, 1.1, 1.1, 1.1, 1.1,
1.2, 1.2, 1.2, 1.2, 1.2, 1.3, 1.3, 1.3, 1.3, 1.3, 1.4, 1.4, 1.4, 1.4, 1.4, 1.4, 1.4, 1.4, 1.3, 1.3,
1.3, 1.3, 1.3, 1.3, 1.2, 1.2, 1.2, 1.2, 1.3, 1.3, 1.4, 1.4, 1.4, 1.4, 1.4, 1.4, 1.4, 1.4, 1.4, 1.3,
1.3, 1.2, 1.2, 1.2, 1.1, 1.1, 1.0, 1.0, 1.0, 1.0, 0.9, 0.9, 1.0, 1.0, 1.1, 1.0, 0.8, 0.8, 0.7, 0.7,
0.6, 0.5, 0.6, 0.5, 0.5, 0.4, 0.4, 0.3, 0.3, 0.3, 0.2, 0.2, 0.2, 0.1, 0.1, 0.1, 0.1, 0.0, 0.0, 0.0,
0.0, 0.0]
```

```
def distribute_ages(population_size):
    # The population is distributed according to the percentages in DATA
    ages = []
    for age, percentage in enumerate(DATA):
        ages.extend([age] * round(population_size * percentage / 100))

    # Add the remainder to the modal age
    ages.extend([DATA.index(max(DATA))] * (population_size - len(ages)))

    return sorted(ages)
```

Appendix 3 – Jobs

```
# Teachers
# Source - https://explore-education-statistics.service.gov.uk/find-statistics/school-workforce-in-england
# 18.0 students per teacher. We have exactly 36 students in the town, so 2 teachers.
# Hours - 8:30 to 16:30
# 2 Teachers

# Retirement Home Staff
# Source - https://lottie.org/care-guides/the-number-of-uk-care-home-residents/
# England - 1.65 staff per resident
# Wales - 1.24 staff per resident
# Average weighted for population of England and Wales - About 1.5 staff per resident
# 10 in the reitrement home, so 15 staff
# Hours - 8:00 to 16:00
# 15 Retirement Home Staff

# Bar Staff
# 3 Workers - 16:00 to 23:00

# Shop Staff
# 5 Workers - 9:00 to 17:00

# Sports Centre Staff
# 2 Workers - 8:00 to 16:00

# Restaurant Staff
# 7 Workers - 11:00 to 20:30

# Theatre Staff
# Not staffed as shows will be infrequently open

# Club Staff
# 2 Workers - 20:00 to 02:00 (next day)

# Unemployed
#https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/timeseries/lf24/lms - April - June 2020 - 75.7%
# We will round this to 85% of adults, as this data takes into account 16-17 year olds, who do not work in our simulation.
# 15% of 99 adults ~= 15 adults
# The remaining 48 adults will be work in workspaces, seperated from the general population.

# Workspace 0 - 6:00 to 14:00  - Construction Site          - 10 adults
# Workspace 1 - 8:30 to 16:30  - Hi-tech Engineering Office -  3 adults
# Workspace 2 - 9:00 to 17:00  - Traditional Office Job     - 16 adults
# Workspace 3 - 9:00 to 18:00  - Administration Job         -  8 adults
# Workspace 4 - 10:00 to 19:00 - Warehouse                  -  8 adults
# Workspace 5 - 12:00 to 20:30 - Low interaction work       -  3 adults
```

Appendix 4 – Model source code link

In line with AQA EPQ Malpractice regulations, the repository is privated, and inaccessible.
The code was written **without** coping or lifting any code from 3Blue1Brown's code: they are written with slightly different libraries, and a full copy of the code can be provided to AQA if needed.
https://github.com/FinleyCooper/EPQ