



University of Southern Denmark

How fair were COVID-19 restriction decisions?

A data-driven investigation of England using the dominance-based rough sets approach

Abel, Edward; Siraj, Sajid

Published in:
arxiv.org

DOI:
10.48550/arXiv.2211.00056

Publication date:
2022

Document version:
Submitted manuscript

Citation for pulished version (APA):

Abel, E., & Siraj, S. (2022). How fair were COVID-19 restriction decisions? A data-driven investigation of England using the dominance-based rough sets approach. Manuscript in preparation.
<https://doi.org/10.48550/arXiv.2211.00056>

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use

This work is brought to you by the University of Southern Denmark.
Unless otherwise specified it has been shared according to the terms for self-archiving.
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.
Please direct all enquiries to puresupport@bib.sdu.dk

How fair were COVID-19 restriction decisions? A data-driven investigation of England using the dominance-based rough sets approach

Edward Abel^{a,*}, Sajid Siraj^{b,c}

^aUniversity of Southern Denmark, Denmark

^bCentre for Decision Research, Leeds University Business School, Leeds, UK

^cCOMSATS University Islamabad, Wah Campus, Pakistan

Abstract

During the COVID-19 pandemic, several countries have taken the approach of tiered restrictions which has remained a point of debate due to a lack of transparency. Using the dominance-based rough set approach, we identify patterns in the COVID-19 data pertaining to the UK government's tiered restrictions allocation system. These insights from the analysis are translated into "if-then" type rules, which can easily be interpreted by policy makers. The differences in the rules extracted from different geographical areas suggest inconsistencies in the allocations of tiers in these areas. We found that the differences delineated an overall north south divide in England, however, this divide was driven mostly by London. Based on our analysis, we demonstrate the usefulness of the dominance-based rough sets approach for investigating the fairness and explainability of decision making regarding COVID-19 restrictions. The proposed approach and analysis could provide a more transparent approach to localised public health restrictions, which can help ensure greater conformity to the public safety rules.

Keywords: multi-criteria data analysis; dominance-based rough sets; COVID-19; data wrangling; data-driven decision making

1. Introduction

The pandemic caused by Coronavirus disease 2019 (COVID-19) has been affecting lives around the globe. The issue of public health and safety, looking to dampen the impacts of COVID-19, forces policy makers to restrict a lot of business and social activities involving physical interactions. However, a prolonged complete lockdown becomes infeasible due to its adverse economic and social impacts, for example, leading to supply chain problems, high unemployment rates, and increases in mental health issues ([Nikolopoulos](#)

*Room 41.58, Universitetsparken 1, University of Southern Denmark, Kolding, Denmark
Email addresses: abel@sdu.dk (Edward Abel), s.siraj@leeds.ac.uk (Sajid Siraj)

et al., 2021; Bassiouni et al., 2023). This makes the issue a multi-objective problem where the safety-related objectives invariably conflict with the economic objectives. Therefore, several governments tried to address this issue by introducing region-level restrictions based on the health and safety risk faced by each region. For example, the UK government announced a tiered restriction system in September 2020 (Department of Health and Social Care and GOV.UK, 2020), as opposed to its previous country-wide restrictions policy. Although this appeared to be a more sensible approach than applying a nation-wide lockdown, some local authorities found this unsatisfactory, due to a lack of sufficient transparency in assigning tiers. Although the government published the set of criteria utilised to determine tier assignments, they did not publish the framework or rules that were used for actual assignment of tiers. As a result, the allocation of tiers during the pandemic has been questioned due to a lack of transparency, leading to accusations of unfair decision making. This study investigates if and how regional inequalities were reflected in the assignment of tiered restrictions during the second wave of the COVID-19 pandemic. Our paper aims to demonstrate the use of a novel Dominance-based Rough Sets Approach (DRSA) to investigate the potential inequalities in policies regarding tiered restrictions.

In this paper, we investigate the potentially different treatment for different geographical areas, specifically focusing on the North-South divide in England. For this purpose, we propose the use of DRSA for understanding the underlying rules that were applied in different geographical areas, and for assessing the level of inconsistencies in assigning those tiered restrictions.

After acquiring and pre-processing data from various disparate sources, we extracted distinct if-then rule sets for different geographical areas and compared these rule sets to identify any inconsistencies. Our analysis found a north south divide, however, the divide was driven mostly by London.

The structure of the rest of the paper is as follows. Section 2 explores literature starting on the tiered restrictions and the use of multi-criteria analysis for this purpose. The usefulness of DRSA is demonstrated with an illustrative example in Section 3. Section 4 then explains our methodology for collecting, preparing and processing data, and performing analysis using DRSA. The findings and analyses are then presented in Section 5, and Section 6 concludes.

2. Background

The UK government applied country-wide lockdown restrictions when the first wave of COVID-19 pandemic hit the nation in March 2020. However, as the data collection for COVID-related metrics matured, disparities within different geographic areas became more visible. It could be sub-optimal to keep the same

level of restrictions to the whole country when there are rising numbers of COVID-19 positive cases (hereafter referred as 'cases' only) only in a specific local area. Hence, when facing the second wave, the UK government announced a tiered restriction system in September 2020 ([Department of Health and Social Care and GOV.UK, 2020](#)), as opposed to a country wide restrictions policy. The tiered approach may mitigate the severity of the impacts of restrictions on economic activities, such as disruptions in the critical supply chains ([Bassiouni et al., 2023](#)).

By monitoring certain measurable indicators, decisions can be made to move areas up a tier (if they are not improving) or down (if the trajectory improves). We discuss this in detail below, in order to make an argument that this is a multi-criteria decision making problem.

2.1. The tiered system

The tiered system, set out in the UK Government's 2020 Winter Plan ([GOV.UK, 2020a](#)), facilitated a more systematic and data driven approach to decision making. They proposed a set of factors (criteria) to determine what level of restrictions (Tier) should be imposed on different geographical areas in England. The restriction were tiered from Tier-1 to Tier-4 ¹, where Tier-1 was the most relaxed set of restriction while Tier-4 represented the most constrained level of restrictions.

The UK government chose a set of five criteria to allocate tiered restrictions. The rationale for the choice was given as '*The indicators have been designed to give the government a picture of what is happening with the virus in any area so that suitable action can be taken.*' ([DHSC and Hancock, 2020](#)). The set of criteria, to determine which Tier from 1 to 4 an area should be placed in, were (C1) case detection rate in all age groups, (C2) case detection in people aged 60 or above, (C3) how quickly case rates were rising or falling, (C4) ratio of positive cases in the general population, (C5) pressure on the healthcare service. In addition to these five criteria, further consideration pertaining to the local context and exceptional circumstances could also be considered, such as a local but contained outbreak. The set of five criteria are defined and discussed below.

2.1.1. Case detection rate in all age groups (C1)

This criterion gives a measure of the number of cases within a given 24-hour time frame for a given geographical area. The cases were detected as positive based on the tests recorded by the UK health system. From this criteria, an indication of how many people are catching COVID-19 in an area can be

¹Initially only Tiers 1-3 were used and then Tier 4 came into effect later

gleamed. However, the measure is only based on tests that are chosen to be taken, therefore, many cases may go undetected, and differences between areas' residents reluctance to go for tests could impact the measure.

2.1.2. Case detection rate in the over 60 year olds (C2)

This criterion gives a measure of the number of cases within a given 24-hour time frame (for a given geographical area) where the patient was over sixty years old. Due to the risk of serious illness from COVID-19 increasing with age this criterion is able to differentiate better the severity of the outbreak in an area than just a overall number of cases, which cannot determine if, the majority of cases are over 60 and thus more serious than if the majority of cases are under 60. Like C1, this criterion is only based on tests that are chosen to be taken and so this self selecting sample group could impact the accuracy of the measure.

2.1.3. The rate at which cases are rising or falling (C3)

This criterion gives a measure of the rate of change of the number of cases between one 24-hour time frame and the next. From this an indication of whether the number of cases is growing or receding can be determined. As this criterion denotes a rate of change, and not an absolute values, it can be the case that high volatility can be present when the number of cases are small. Moreover, like C1 and C2, this criterion is also based on tests that are chosen to be taken so again represents a self selecting sample group.

2.1.4. Positivity rate (C4)

The positivity rate criterion is a measure of the number of positive cases that are detected as a percentage of all the tests taken within a given 24-hour time frame. From this an indication of the general prevalence of COVID-19 can be extrapolated, however, it is calculated only from tests that are chosen to be taken, and potentially those that are more likely to have COVID-19 symptoms may be more likely to look to confirm this with a test.

2.1.5. Pressure on the NHS (C5)

This criterion gives an indication of the pressure that the cases are having on the NHS infrastructure for a given geographical area. This is an important consideration as if the pressure were to become so high that the NHS infrastructure becomes overwhelmed then, its ability to tackle the cases that result in hospitalisation would be severally hampered, which in turn could have highly negative impacts on its ability to prevent some cases from ultimately resulting in deaths. This criterion is based on numerical information

regarding hospital admissions, so unlike other criterion, such as those taking COVID-19 tests, is not based on voluntary participation. However, it only considers situations requiring health service interventions.

2.2. Were these tiered restrictions fair?

The government claimed that the tiered system would result in greater transparency, however, people questioned the integrity of the system from the start. For example, Manchester was forced to enter into Tier-3 without sufficient evidence to appease local policy makers and the local community.

To investigate the issue of fairness in assigning tiered restrictions, it is important to first understand the concept of fairness. When talking of regional fairness and equality, most research has focused on the distribution of resources. For example, fair distribution of food ([Alkaabneh et al., 2021](#); [Busch and Spiller, 2016](#)), assigning public health resources ([Tavana et al., 2021](#); [Norheim, 2016](#)), or public transport ([Gaile, 1977](#)) between different regions. In other cases, this concept is discussed for allocating budgets (finances) in a fair and equitable manner ([Shankar and Shah, 2003](#)). One can argue that budget is also a type of resource distribution where the resource is a limited monetary value. Both issues involve a indirect (or indirect) competition as different regions are competing against each other for obtaining (or maximising) their share of finite resources. However, the idea of assigning different levels of restrictions does not fit into these two types, as the assignment of tier to any region is independent of how other regions are performing (or which state they are in). Thus the concept of fairness in a tiered allocation system is potentially different from the fairness in distribution of resources, where allocating more resources to one region implies the other regions will get less. Therefore, assigning heavy restrictions in one region has no dependence on the restrictions assigned to other regions.

Although assigning restrictions to one region is independent of other regions, people have a natural tendency to compare the restrictions imposed on them with those in other regions ([Alicke and Zell, 2008](#)). Any preferential treatment to one region over others might not be welcomed, and might end up in some form of protest. This protest can take a form of doing (or not doing) actions that lead to breaking the rules ([Williams et al., 2020](#); [Carthey et al., 2011](#)). For example, in the case of allocating tiers, people in higher tiers still have the ability to break the rules, and act as if they were in a lower tier. Therefore, when they consider it unfair, it is more likely to see people breaking the rules, which makes these decisions critical for regional peace and stability.

2.3. North South Divide in England

In England, the North-South divide is a term that refers to the socioeconomic differences between Southern and Northern parts of England ([The Economist, 2012](#)). A recent report by the Institute for Public Policy Research ([Webb et al., 2022](#)) states that the South of England consists of one-third of the UK population yet accounts for 45% of its economy and 42% of its wealth. Considering these numbers, it is important to investigate the possibility of this North-South divide penetrating in the COVID-19 related policies as well.

Before discussing this further, it is important to first define the boundaries separating the North and South. In terms of creating boundaries between North and South, a widely accepted belief is that the North consists of the regions of the North East, North West, Yorkshire and The Humber, East Midlands, and West Midlands ². On the other hand, the South consists of the East of England, London, South East, and South West. Although London is considered part of South, we are interested in exploring how it has been argued that London has its own unique dynamics due to its different socioeconomic demographics. Therefore, for research purpose, we will also analyse London separately due to its unique data properties (explained in Subsection 5.1).

For the purpose of our study, we will therefore consider three geographical area categories of i) North, ii) South Sans London, and iii) London. These three areas can be seen geographically in Figure 1 where each of the nine regions have been coloured green, red or blue. The green regions constitute North while the blue regions constitute South without London, and London is shown in red colour.

Although the assignment of tiers is not a zero-sum game - where one person's gain is another person's loss - any inconsistency in the assignments can have serious implications as people would compare their level of freedom (or level of safety) to the other regions, and these comparisons might dissuade some of them from sticking to the rules ([Moss, 2020](#)).

A recent review of the use of data analytics in COVID-19 concluded that it has been successfully employed in the healthcare sector ([Khan et al., 2021](#)). We argue that the use of data analytical techniques can also help identify any anomalies or inconsistencies in the assignments of the tiered restrictions. In this context, the use of dominance-based rough sets approach can be useful. Before we investigate its usefulness in investigating fairness, let us briefly introduce this approach in the next section.

²However, in some studies, the Midlands have also been considered a separate geographical entity from the North and the South ([Smith, 2018](#))

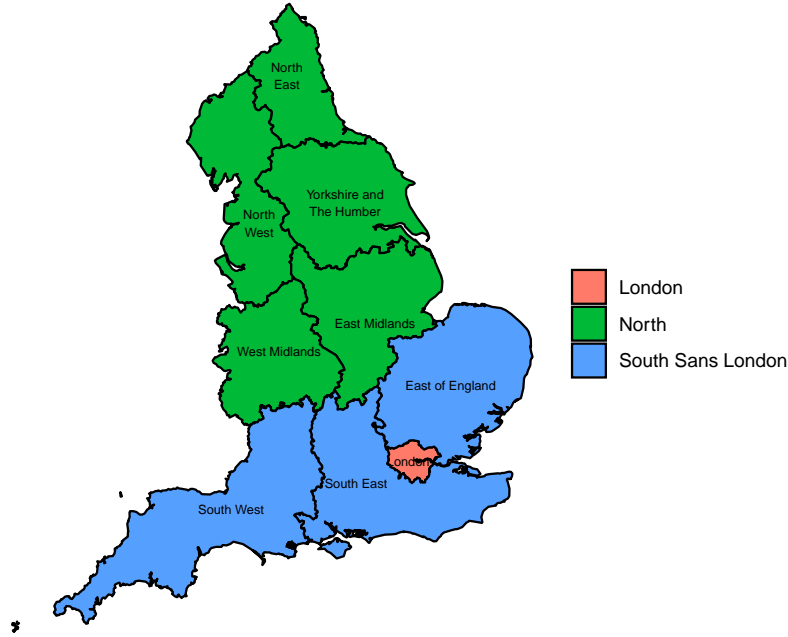


Figure 1: England Regions

3. Investigating COVID restrictions using DRSA

Dominance-based Rough Set Approach (DRSA) is a well-known method for multi-criteria classification, which is used to extract understandable IF-THEN decision rules from analysing historical data (Greco et al., 2001). DRSA has been applied to numerous applications from detecting frauds (Błaszczyszński et al., 2021) to analysing service quality (Liou et al., 2011). Let us demonstrate the usefulness of DRSA in analysing COVID-related data using an illustrative example below.

Table 1 provides synthetic data with ten observations, where each observation has three input readings relating to the number of cases, rate of change (in the number of cases from one day to the next), and the positivity rate. The fourth value can be considered as an output showing the Tier level allocated to that observation. In the DRSA literature, each observation is sometimes referred to as object; and the inputs are sometimes referred to as criteria attributes, and the output referred to as decision attribute.

For example, in Table 1 the second observation (x_2) shows that there were 92 (positive) cases, while the rate of change was 2.45, and the positivity rate was 7.89%. When comparing this observation with the first observation (x_1), we can see that all three input values of x_2 are lower than their respective values for x_1 . Therefore, we can say that x_2 dominates x_1 . In this way, we can compare the inputs of each observation with every other observation to determine its dominance. Table 2 summarises this dominance relationships for all

Table 1: Illustrative example to demonstrate the use of DRSA for COVID-related data

Observation	Number of Cases	Rate of change	Positivity Rate	Tier
x_1	195	2.48	8.05	3
x_2	92	2.45	7.89	2
x_3	237	-2.74	8.94	2
x_4	515	2.82	1.43	3
x_5	528	7.54	5.3	3
x_6	434	1.65	5.41	2
x_7	143	-3.15	8.01	1
x_8	75	3.2	5.25	2
x_9	269	2.33	1.71	1
x_{10}	131	3.28	1.03	1

Table 2: Summary of dominating and dominated observations

Observation	Dominating	Dominated
x_1	1	1,2,7
x_2	1, 2	2
x_3	3	3, 7
x_4	4, 5	4
x_5	5	4, 5, 8, 9, 10
x_6	6	6
x_7	1, 3, 7	7
x_8	5, 8	8
x_9	5, 9	9
x_{10}	5, 10	10

observations. For example, the fifth observation (x_5) is clearly dominated by four other observations, those of x_4 , x_8 , x_9 , x_{10} . Note that the table also includes x_5 itself as, by definition, the dominance relationship also includes equality.

After processing all the dominance relationships in Table 2, we also need to process the outputs, which are essentially the Tier values assigned to each observation. As we know that tiers are ordered categorical values, we can group these tiers into unions by enumerating all possible combinations. This is summarised in Table 3 where we have four unions of "at most T1", "at most T2", "at least T2" and "at least T3". The unions involving "at most" are also known as *downward* unions, and those involving "at least" are known as *upward* unions.³

Now that the inputs and outputs are processed separately, the next step would be to somehow link the two tables together. This is usually done by inducing rules using rough-set based algorithms like DOMLEM (Greco et al., 2000) or VC-DOMLEM (Błaszczyszński et al., 2011). The set of rules induced using DOMLEM

³Please note that the unions of "at least T1" do not make sense as all observations will be at least T1.

Table 3: The unions of classes for the illustrative example

Union	Observations that are part of the union	Total
at most T1	7, 9, 10	3
at most T2	2, 3, 6, 7, 8, 9, 10	7
at least T2	1, 2, 3, 4, 5, 6, 8	7
at least T3	1, 4, 5	3

Table 4: Illustrative example to demonstrate the use of DRSA for COVID-related data

No.	Antecedent	Consequent	Support	Strength
1	(Number of Cases ≤ 269) and (Positivity Rate ≤ 1.71)	at most T1	2	66.67
2	(Rate of Change ≤ 2.45)	at most T2	5	71.43
3	(Number of Cases ≤ 434) and (Positivity Rate ≤ 5.41)	at most T2	4	57.14
4	(Number of Cases ≤ 131)	at most T2	3	42.86
5	(Number of Cases ≥ 195) and (Rate of Change ≥ 2.48)	at least T3	3	100.00
6	(Number of Cases ≥ 515)	at least T3	2	66.67
7	(Rate of Change ≥ 2.82) and (Positivity Rate ≥ 1.43)	at least T2	3	42.86
8	(Number of Cases ≥ 434)	at least T2	3	42.86
9	(Rate of Change ≥ 1.65) and (Positivity Rate ≥ 5.41)	at least T2	3	42.86

for this example are shown in Table 4, where each rule has an antecedent and a consequent. The antecedent is the set of conditions that are required for this rule, and the consequent is the resulting union.

Out of the nine rules extracted for this example, one can see that Rules 2, 4, 6 and 8 (see grey shaded rules) have only one condition to satisfy, and hence we can say that the rule length is 1. Consequently, the remaining five rules are of rule length 2 as all of them need two conditions to be satisfied.

The first four rules in Table 4 have *consequents* involving "at most" while the remaining six rules involve "at least" in their *consequents*. The rules related to "at most" are termed as downward rules, whilst the rules related to "at least" are termed as upward rules. The last two columns in Table 4 show the support and strength for each rule induced from the data. The support refers to the number of observations that adhere to this rule, while the strength refers to the ratio of observations that met the rule against the observations that only met the first part of the rule (i.e. *antecedent*). For example, for the rule in row 1 of 4, there are three observations under the *antecedent*, of which 2 adhere to the rule (hence a support of 2). The strength for the rule is therefore, $3/2=66.67\%$.

Some recent studies on COVID-19 demonstrate the benefits of using DRSA for data analysis and predictive modelling. For example, [Figueiredo et al. \(2022\)](#) considered a number of criteria and risk factors for COVID-19 and proposed the use of DRSA to identify areas vulnerable to COVID-19 infections in Brazil.

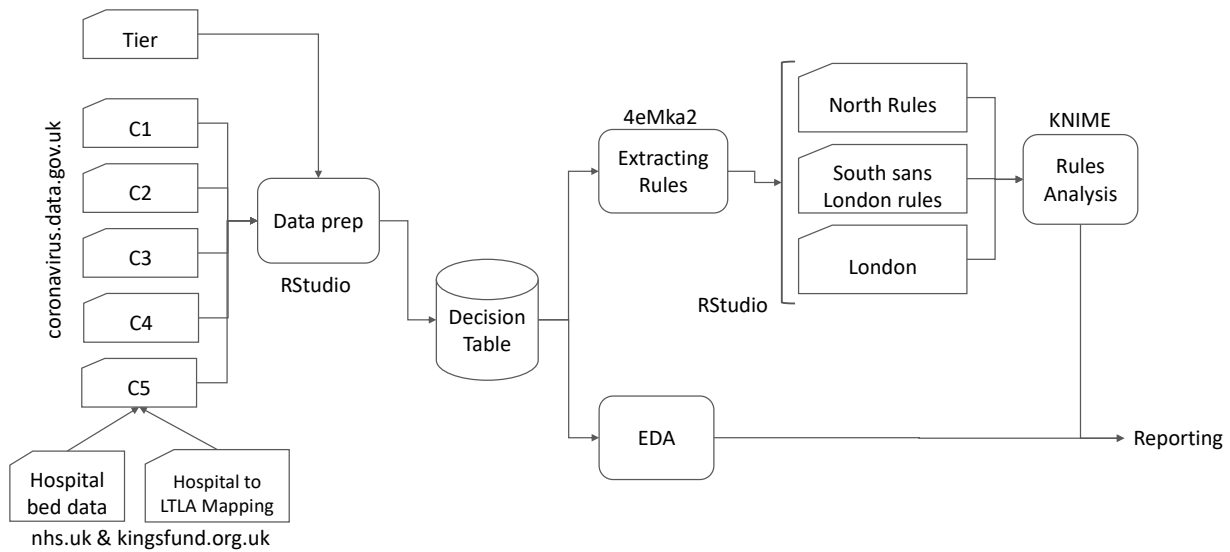


Figure 2: The Data Extraction, pre-processing and analysis pipeline of our approach

Similarly, [Bhaskar et al. \(2021\)](#) proposed the use of rough sets to predict symptomatic cases of COVID-19 patients. In these and other existing studies, the focus remains on extracting a single set of rules from historical data. However, in this research, we propose a novel use of this approach, that is, to extract separate sets of rules for different geographical segments. Although these sets of rules might be useful in a number of ways, we demonstrate how we can compare these rules for investigating fairness across different geographical areas - explained in the next section.

4. Methodology

The UK government’s tiered allocation assignments were based on data pertaining to the prevalence and risk posed by COVID-19 in different parts of England. In the illustrative example earlier, a set of rules was derived from a set of observations. Comparisons between different geographical areas can be performed by performing this process of rules derivation for different areas of England, utilising only the observations specific for each area, to generate a separate set of rules for each area.

In this way, from the extracted sets of rules for each geographical area under comparison, we have the opportunity to assess the overall fairness of the tiered system. To explore the fairness of the tiered approach in relation to the North-South divide, we can look to generate three separate rule sets from the observations from (i) the North, (ii) the South sans London, and (iii) London, as defined earlier and shown in [Figure 1](#).

The process pipeline of collecting and analysing data is shown in [Figure 2](#); which includes data acquisition,

Criterion	Description	Data source(s)
C1	Number of Daily Cases (all ages)	coronavirus.data.gov.uk
C2	Number of Daily Cases in over 60s	coronavirus.data.gov.uk
C3	Rate of change of daily cases	Derived from C1 data
C4	Positivity rate	api.coronavirus.data.gov.uk
C5	Pressure on the NHS	1) api.coronavirus.data.gov.uk 2) www.kingsfund.org.uk 3) github.com/epiforecasts
Decision Variable	Tier Value	Parliament Legislation documentation

Table 5: Data sources for government-defined criteria set

pre-processing, and the extraction and analysis of DRSA rules. To explore such an approach requires the collection of observation data relating to the set of government criteria over the duration of the tiered allocation system. For our approach data was collected from various different disparate sources. The following section explains the data collection in detail.

4.1. Data acquisition

For the different areas of England, data was obtained pertaining to the UK government’s set of five criteria discussed earlier (DHSC and Hancock, 2020), along with the allocated tier values. Table 5 provides the list of data sources used to collect data on these five criteria.

England is geographically broken up into 9 Regions which are further broken down into a set of Lower Tier Local Authority (LTLA) areas (ONS, 2016). Data relating to COVID-19 at the LTLA level of granularity can be collected. Therefore, we curate a data-set consisting of Tier value and criteria calculations, for each LTLA, for each day that the Tiered system was in place. We briefly outline the data acquisition process for each criterion below.

4.1.1. C1: Number of Daily Cases

The website www.coronavirus.data.gov.uk is a UK government website that provides official data relating to coronavirus (COVID-19). The data denoting the number of positive cases which were reported in each LTLA region every day can be obtained from the site ⁴. Although the data is provided by age demographics, the overall number of cases were derived by totalling the cases across the set of age ranges.

As a common practice (Katris, 2021), we calculated a 7-day rolling average to alleviate discrepancies in the reporting velocity at different days of the week. The effects of this process are shown in Figure 3, where,

⁴From the site’s Supplementary downloads

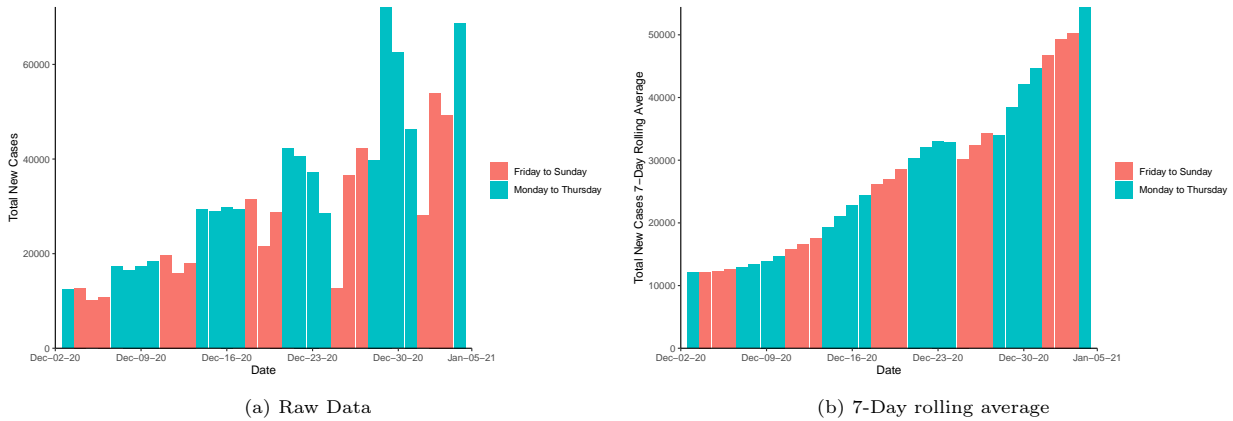


Figure 3: Raw data and 7-day rolling average data

the figure on the left clearly shows that the number of reported cases around weekends is invariably lower than the cases on other days. The figure on the right shows the rolling average where this issue has been addressed.

4.1.2. C2: Number of Daily Cases in those aged 60 plus

Recall that the the number of cases collected for C1⁵ are broken down by age demographic ranges, therefore, we derived C2 values as the sum of age demographic ranges 60 and over (i.e. for each LTLA for each day). As with C1, we then calculated a 7-day rolling average for C2 as well.

4.1.3. C3: Rate of Change in Number of Daily Cases

The rates of change in the number of cases (from one day to the next) can be determined from the data obtained for C1. A 7-day rolling average was calculated for these rate of change values for each LTLA region for each day.

4.1.4. C4: Positivity rate

The positivity rate of cases denotes the percentage of tests taken over a given time period that are returning positive results. The positivity rate for each LTLA region for each day can be sought. Such data can be retrieved via the UK government’s COVID-19 API⁶ Through utilising the API making a set of call requests, for each LTLA region, we obtained the positivity rate, for each LTLA region for each day, which again we processed into 7-day rolling average values.

⁵from the coronavirus.data.gov.uk site

⁶api.coronavirus.data.gov.uk - providing API calls for a range of COVID-19 related metrics and levels of geographical granularity.

4.1.5. C5: Pressure on the National Health Service (NHS)

The criterion of pressure on the NHS looks to measure the amount of pressure being placed on the local medical services, in terms of their ability to be able to cope with the COVID-19 cases prevalent within a given area. The UK government utilised this criterion in their tiers allocation system, however, the pressure on NHS cannot be quantified directly at LTLA level. To estimate the pressure, we relied on the number of beds occupied by COVID-19 patients at NHS Trusts level. The NHS service is broken geographically across England into a number of Trusts, 223 in total⁷ (which each operate as a somewhat automated unit).

Due to lack of government transparency in quantifying the pressure on NHS, the Trust level values were collected and mapped to LTLA level to quantify the pressure on the NHS, as there was a need to align the Trust level data with the LTLA level data. This mapping (performed in R) is explained below.

Using the UK government's COVID-19 API provides the number of beds occupied by COVID-19 cases (for each trust for a given day). NHS trusts across England vary significantly in size (and thus levels of capacity) therefore, comparisons between such values on their own may be misleading. For example, one trust may have double the number of beds occupied with COVID-19 patients than another but still be under less pressure due to being three times the size. The government documentation regarding the tiers system remains opaque in terms of how such information was exactly utilised to help make decisions. However, we can utilise various other publicly available data to derive measures of the relative pressure on the NHS.

Each trust has multiple hospital sites for which data relating to the total number of available beds for each trust can be sought⁸. We already know the actual number of beds occupied by COVID-19 patients so therefore, we can estimate a representative ratio of occupied beds in relation to the overall number of beds in each trust.

Recall these values are available for each trust, however, to perform these analysis, we need all values at the LTLA level. A trust may serve more than one LTLA region, and an LTLA region might be served by multiple trusts. Therefore, in this way, there exists a many to many mapping between trusts and LTLAs.

One way to resolve this issue is to calculate probabilistic estimates through mapping the information at trust level to LTLA level. To achieve this we utilised the Trust to local authority mapping provided by Epiforecasts⁹. This package provides probabilistic mapping estimates that can be used to estimate COVID-19 hospital admissions at an LTLA level. From utilising this mapping on our data we obtained a ratio based pressure value for each LTLA region for every day, and thus at the same level of granularity that we have

⁷www.kingsfund.org.uk/audio-video/key-facts-figures-nhs

⁸www.kingsfund.org.uk

⁹Available as a developer package on GitHub at - github.com/epiforecasts/covid19.nhs.data

for the other four criteria.

4.1.6. Decision Variable: Tiers Data

Date pertaining to tier values (1,2,3 or 4) that different parts of England were placed under, and subsequent changes, were announced in the houses of Parliament and published in official legislation documentation by the government (GOV.UK, 2020b). From this we were able to define the Tier that each LTLA region was in for every day that the tiered system was in place.

4.2. Data processing and integration

From the acquisition and format alignment processing of data from these various sources making up the set of 5 criteria we then integrated the data to create an overall data-set. In the data-set each row (observation) denotes the data for the set of 5 criteria and the tier pertaining to an LTLA and Date pair, with additional information added to each observation regarding its membership of either Northern England, Southern England sans London, or London. The data-set consists of 10827 observations. ¹⁰

The data-set was then utilised to gain understanding about the allocation of tiers, first with an exploratory data analysis, and then using the DRSA rules analysis approach. These two approaches are described in the next subsections below.

4.3. Exploratory Data Analysis

For better data understanding, we performed some statistical analysis of the data-set. This included analysing the spread of decision variable values, across the range of dates, and between the different geographical areas. We then analysed the relationships between and within the criteria set and tier values allocated. The results for exploratory data analysis are discussed in Section 5.1.

4.4. DRSA Rule Extraction

To explore the fairness and validity of the tiers allocation system via DRSA, rule extraction and comparison of rule sets was performed.

The overall data-set was divided into three separate files representing observations pertaining to the North, the South sans London and London. These observation files were then processed ¹¹ to generate information system files (.isf) readable by 4eMka2¹². These files were then fed into 4emka2 that generated

¹⁰The enriched data-set for this study is publicly available on github here: https://github.com/prioritization/DRSA_Covid19

¹¹using R scripts to automate the process

¹²4eMka2 is a tool developed by the Laboratory of Intelligent Decision that implements the DRSA method Support System of the Institute of Computing science, Poznan University of Technology, idss.cs.put.poznan.pl

three output files, providing us the three separate sets of extracted rules files (.rls). This process is also summarised in Figure 2 (see top-right of the figure). Regarding rule extraction parameters, initially only the rules with a minimum rule strength of 25% were saved, with a maximum rule length of up to 5 (i.e. all of the criteria). The output .rls files were then parsed and tabulated (using R scripts for automation) for further processing. Then, using the mining tool KNIME ¹³ the separate rules were analysed and compared.

5. Analysis and Results

In this section the results of the exploratory data analysis, and the analysis and discussions of using the DRSA rules analysis approach are described.

5.1. Exploratory data analysis

A set of exploratory data analyses were carried out to gain data understanding before investigating further. These are described and discussed below.

5.1.1. Spread of tiers values

We explored the distribution within the data of Tier values across the time period. Figure 4a shows these distributions for the whole data-set, where it can be seen that a majority of areas were in Tier-2 in early November but moved into Tier-4 by mid December. Tier-1, however, remains infrequent in the whole data-set, with only 0.5% of observations, (the other tier values percentages of observations in the whole data-set are Tier-2, 37.9%, Tier-3, 37.8%, and Tier-4, 25.9%).

Figure 4b shows that the majority of the Northern regions started and stayed in Tier-3 while the Southern regions (sans London) started in Tier-2 (see 4c) and stayed in Tier-2 until towards the end of December. London followed a similar but even more pronounced pattern than the South, as visible in 4d.

The analysis of Figure 4 highlights the imbalance of the distribution of the Tier values across different regions. This imbalance might impact the coverage of DRSA rules that can be generated, for example, rules about Tier-1 might not be easy to derive due to the sparsity of observation data with Tier-1.

Observation 1. *There is an imbalance of the distribution of the Tier values.*

¹³An open-source data analytics, reporting and integration platform - www.knime.org

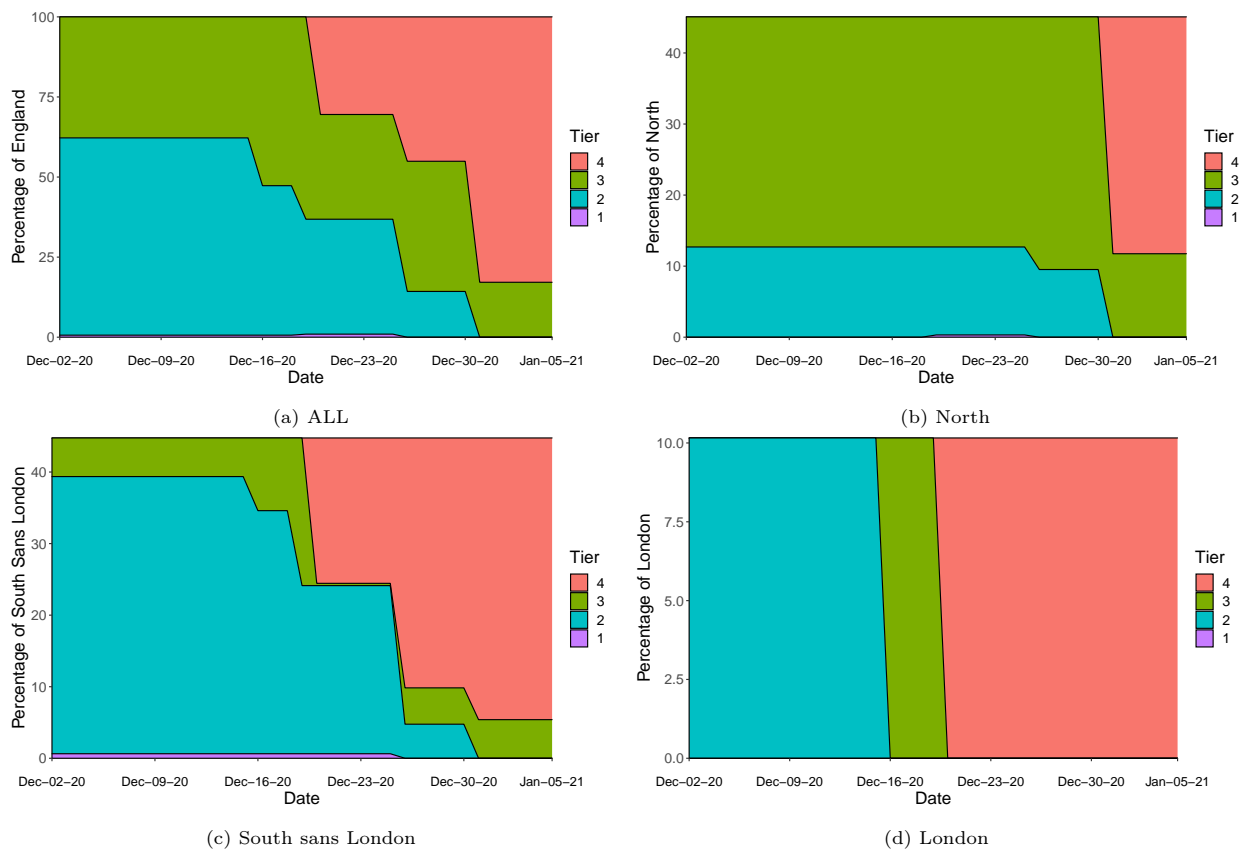


Figure 4: Density Plots of the Distribution of Tiers across the Data Range for different geographic areas

5.1.2. Analysis of relationships between Tier values and criteria values

Next, we analyse potential correlations between the criteria values and Tier values, across the different geographical areas. These correlations are shown in a set of tables provided in Table 6. Looking at Table 6a - showing pairwise correlations among criteria and Tiers within the whole data-set - one can see that there is a strong positive correlation between C1 and C2 (i.e. 0.94). Although this is true for all four tables, the correlation in the data from North (see Table 6b) is slightly stronger than in the data obtained from The South Sans London and London (see Tables 6c and 6d). However, regarding the other criteria the North has significantly lower correlations amongst the set of five criteria (and Tier values) than The South Sans London and London.

Observation 2. *The number of cases in age group 60+ (C2) is highly correlated with the overall number of cases (C1)*

Observation 3. *Data for the North has significantly lower correlations amongst the set of five criteria (and Tier values)*

A further insight from the correlation data is that C3 (Total New Cases Rate of Change) has a weak correlation with the Tier value (as well as to the other criteria). Conversely, C4 (Positively Rate) has the strongest correlation with the Tier value for the overall data (as well as for each of the three geographical breakdowns). This suggests that the 'Total New Cases Rate of Change' has little influence upon Tier values, whilst the 'Positivity Rate' does.

	C1	C2	C3	C4	C5	Tier
C1		0.94	0.01	0.74	0.61	0.49
C2			-0.02	0.74	0.54	0.53
C3				-0.06	-0.06	-0.10
C4					0.56	0.71
C5						0.36
Tier						

(a) Correlations for overall data

	C1	C2	C3	C4	C5	Tier
C1		0.96	0.12	0.52	0.26	0.32
C2			0.11	0.53	0.25	0.34
C3				0.11	-0.03	-0.00
C4					0.25	0.65
C5						0.28
Tier						

(b) Correlations for North region

	C1	C2	C3	C4	C5	Tier
C1		0.93	-0.09	0.80	0.68	0.58
C2			-0.13	0.80	0.62	0.62
C3				-0.18	-0.10	-0.16
C4					0.61	0.75
C5						0.40
Tier						

(c) Correlations for South region Sans London

	C1	C2	C3	C4	C5	Tier
C1		0.93	-0.27	0.78	0.41	0.62
C2			-0.10	0.76	0.44	0.64
C3				-0.16	-0.11	-0.12
C4					0.52	0.76
C5						0.45
Tier						

(d) Correlations for Greater London

Table 6: Correlation table...

These relationships are shown visually through Box-plots in Figure 5. Here, Box-plots are shown for C1,

C3, C4 and C5. C2 is omitted due to the strong positive correlation previously identified between C1 and C2, which results in very similar box-plots for C1 and C2. In each of these plots, the criteria values are grouped according to the Tier values, and a statistical summary is shown for North, South Sans London, and London. For C1, C4 and C5, the Box Plots show that London was invariably assigned to be in a lower Tier for values that would have resulted in higher Tier assignment for the North and the South Sans London. For example, for C1, the median value for London being assigned Tier-3 is 239.43, where as conversely, the median value for North and South being assigned to Tier-4 is 77.14, and 115.50000 respectively. This suggests London was given different treatment, perhaps showing that it was being favoured for economic activities at the expense of public health; suggesting implicit additional criteria exist within the Tiers system.

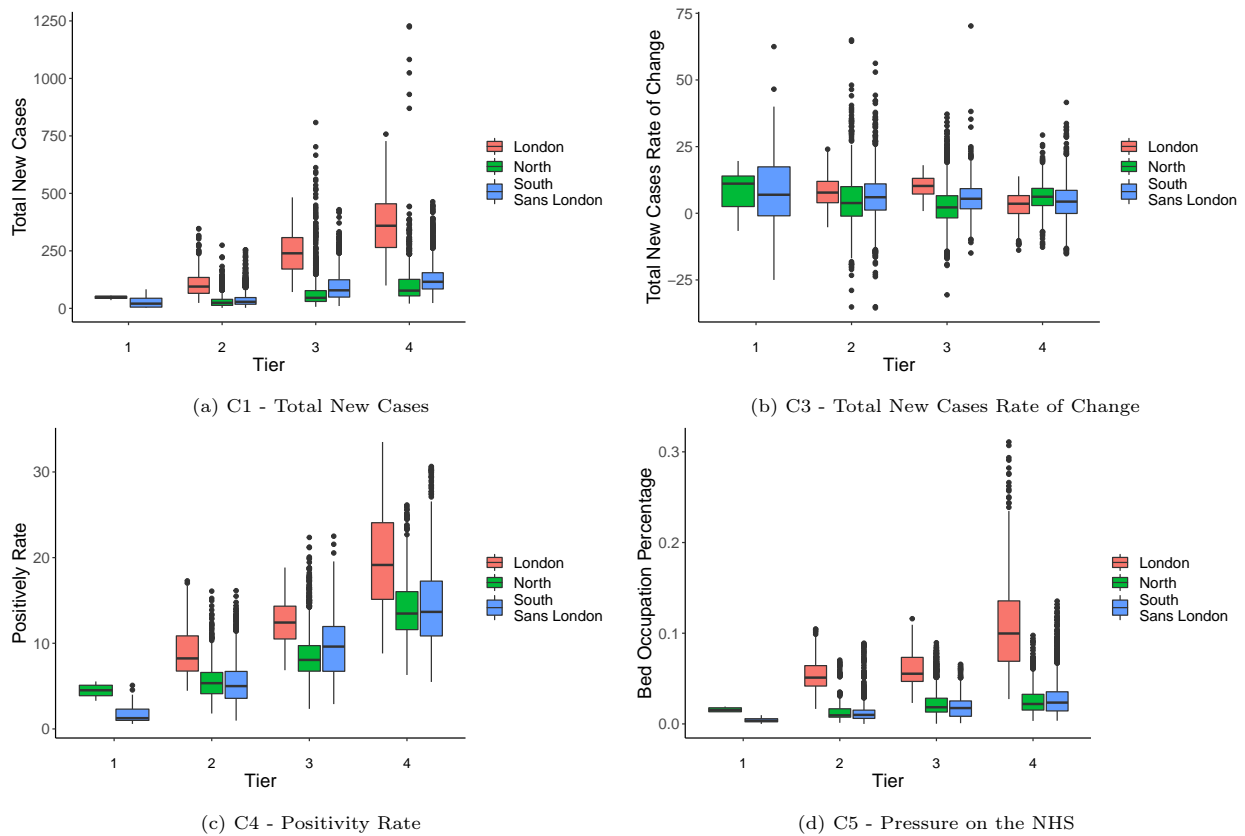


Figure 5: Box Plots

Observation 4. For C1, C4 and C5, London was allocated to be in Tier-2 or Tier-3 for the values that put other geographical regions in Tier-4.

From these box-plots, for C3 (Total New Cases Rate of Change) shown in Figure 5b, there is apparently no relation between the criterion's value and the assigned Tier. Whereas, for C4 (Positivity Rate), a clear positive relationship is visible with the Tier value (albeit at different rates for the different geographical

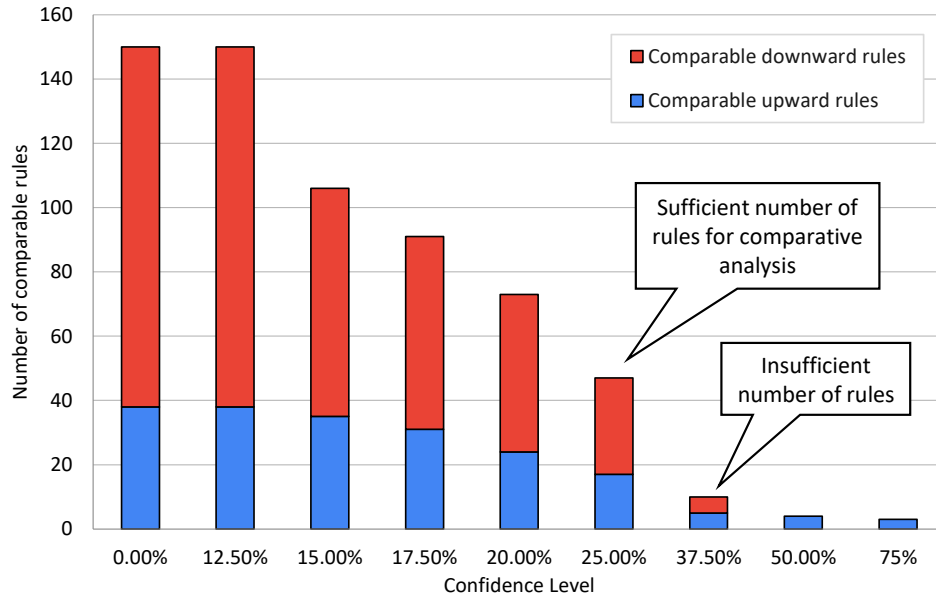


Figure 6: Increasing the confidence threshold reduces the number of possible comparisons

regions). This reiterates what the correlation analysis highlighted, namely that C3 appears to have a weak or no relationship with an assigned Tier value where as C4 does.

Observation 5. *The Rate of change (C3) has a weak or no relation with Tiers allocation, whilst C4 (Positivity Rate) appears to have a strong relation with Tiers.*

So far, each criterion has been investigated individually, without analysing possible interrelationships amongst them and the assigned Tiers. Also, to investigate the fairness, it is more appropriate to investigate the combinations of tiers, for example, investigating geographical area in "at least" Tier-3 or "at most" Tier-2. Such analyses can be carried out using DRSA, as discussed in the next subsection. The DRSA models generated from such analysis can also be used to help us sort unseen data, and in turn, facilitate predictive analysis for evidence-based decision making.

5.2. DRSA Rule analysis and comparison

Although the DRSA has been widely used in practical applications, these applications have focused on a single set of rules extracted from a data-set. On the contrary, in this research, we focus on generating multiple sets of rules created from different data-sets, and then comparing these rule sets for gaining further insights.

The rules were extracted using the 4emka software tool using the rule strength as a filter to include/exclude rules for comparison. Keeping the rules with lower strength may reduce the quality of information for com-

		tier at least 2							tier at least 3			tier at least 4		
		C1	C2	C4	C5	C1C5	C2C5	C3C4	C3C4C5	C4C5	C1C4C5	C2C4C5	C4	C4C5
C1	London	297.14				-					-			
	North	62.04				49					80.35			
	South	-				55.43					130.39			
C2	London		33.71				-							
	North		-				10.71					17.44		
	South		14.19				9.18					20.8		
C3	London													
	North							-1.37	5.44					
	South							1.42	-3.68					
C4	London			23.5										
	North			6.36				5.14	4.64	9.09	9.84	9.43	22.67	19.81
	South			5.12				4.5	2.86	13.64	13.25	11.55	22.63	16.49
C5	London				4.22	-	-							
	North				2.09	1.86	1.86		1.35	2.91	1.87	1.54		3.55
	South				1	0.46	0.46		0.83	3.3	1.56	2.07		4.88

Table 7: Upwards Rules Table

paring different geographical regions. However, on the other hand, filtering out these rules might result in having sparse information which could be insufficient for comparative analysis. This can be seen as a quantity-quality trade-off problem and finding a balance between these two objectives is an important issue to address. Recall the lack of coverage of our data-set shown in Figure 4a. This already constrains our analysis due to a limited number of rules having higher confidence values.

Figure 6 shows the results of an experimental trade-off analysis carried out by changing the threshold level of confidence which can help determine a trade-off between confidence level (i.e. quality) and the number of comparable rules (i.e. quantity). As shown in this figure, the number of comparable rules decrease significantly as we increase the threshold of acceptance for confidence. Therefore, in this case study, we used the minimum rule strength of 25% to keep sufficient information required to compare different regions. However, we believe the selection of this threshold is both 1) context-dependent, where other input data sets might result in alternative thresholds to be chosen, and 2) User dependent, subjective to different user dispositions regarding the trade-offs involved.

Instead of listing all the rules in a traditional if-then format, we display these rules in a table in order to better compare the three regions (with respect to different outcomes). The extracted upward rules are summarised in Table 7, while the downward rules are summarised in Table 8. One may argue that the upward rules are about restricting regions from entering a lower Tier and so could be considered more relevant from health policy perspective, where as downward rules restrict the regions from entering a higher tier and so are more concerned with the economic considerations.

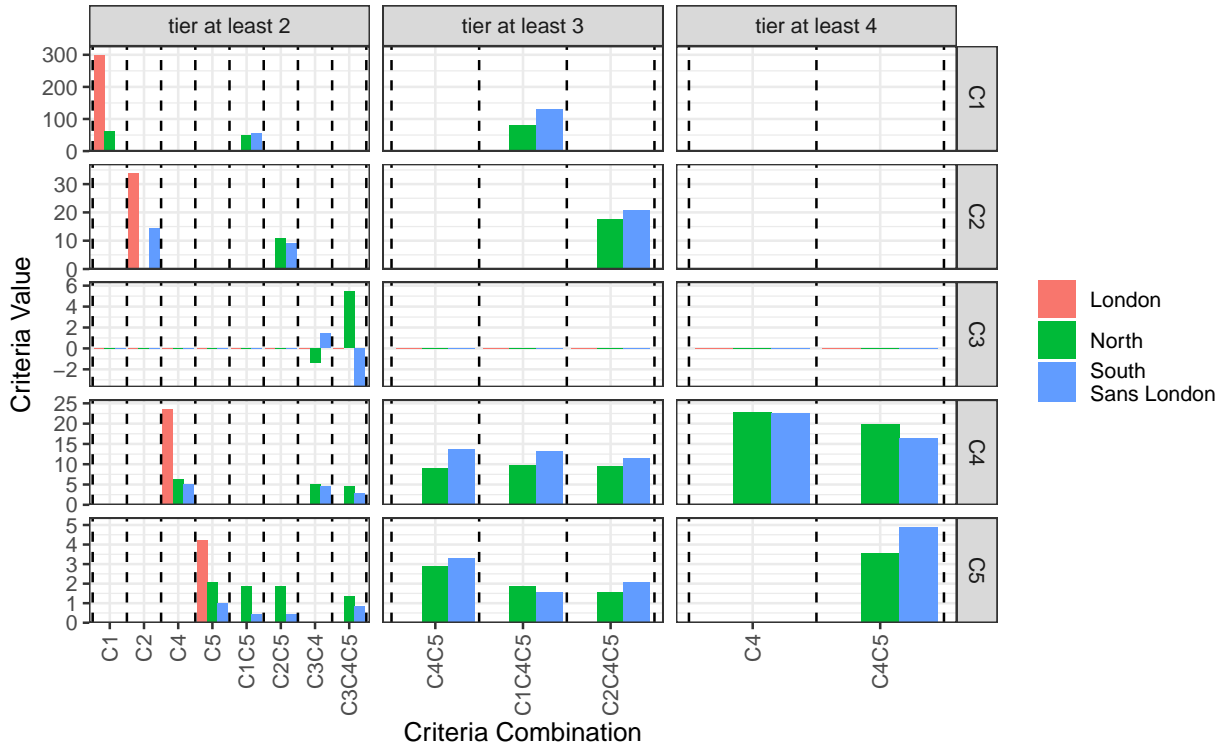


Figure 7: Upward Rules Plot

Taking the first criterion C1 (number of new cases) with consequent "tier at least 2" in the upward rules table (Table 7), the extracted rules suggest that London is at least in Tier 2 for values greater than 297.14, whereas the North is at least in Tier 2 for values greater than 62.04. This implies that the thresholds for London are relatively more relaxed than the North. For example, if the number of new cases is over 62.04 and below 297.14, the extracted rule suggests that North should not be in Tier-1, whilst London might still be in Tier-1. The ratio between these values is 1:4, suggesting London's cases can be almost five fold more before it is treated the same as the North. Note that here there was no such rule extracted for the South to compare with the North and London.

Observation 6. *The thresholds for upward rules give evidence of relaxed rules for London for C1*

Focusing on the second criterion of C2 (number of cases in those aged 60+), DRSA has extracted rules for London and the South, however, no rule was extracted for the North. As observed for C1, the threshold for London (i.e. 33.71) are also more relaxed than the threshold for the South (i.e. 14.19). The threshold for London is more than twice the threshold for the South. In this case, there was no rule extracted for the North.

For the fourth criterion of C4 (Positivity rate), DRSA has extracted rules for all three of the geographical

		tier at most 3													
		C1	C4	C1C2	C1C4	C2C4	C3C4	C4C5	C1C4C5	C2C3C4	C2C4C5	C3C4C5	C1C3C4C5	C2C3C4C5	C1C2C3C4C5
C1	London	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	North	19.5	-	21.14	22.96	-	-	-	31.55	-	-	-	56.93	-	33.51
	South	21.09	-	31.24	42.24	-	-	-	38.29	-	-	-	37.71	-	42.6
C2	London	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	North	-	-	4.29	-	6.51	-	-	-	8.84	8.41	-	-	7.63	6.54
	South	-	-	3.91	-	5.18	-	-	-	4.86	6.23	-	-	7.08	6.52
C3	London	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	North	-	-	-	-	-	-0.08	-	-	5.19	-	6.58	12.64	11.44	11.17
	South	-	-	-	-	-	10.44	-	-	4.42	-	8.82	8.26	8.61	6.78
C4	London	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	North	-	6	-	8.14	7.78	8.36	6.75	8.24	8.2	7.19	9.13	7.94	8.81	8.43
	South	-	5.32	-	5.57	7.28	5.55	6.22	6.18	7.31	6.65	5.72	7.09	6.76	6.95
C5	London	-	-	-	-	-	-	1.3	1.77	-	1.67	1.04	1.48	1.81	2.28
	North	-	-	-	-	-	-	1	1.27	-	1.27	1.6	1.15	1.28	1.63
	South	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 8: Downward Rules Table

regions, facilitating a comparison between all three. Here again, the threshold for London (23.5) is many folds higher than others, whilst there is a negligible difference between the North (6.36) and the South (5.12).

Observation 7. *The thresholds for upward rules give evidence of relaxed rules for London for C4 as London was in Tier-2 when others were in Tier-3*

Considering the fifth criterion (Pressure on NHS), again the threshold for London remains very high (4.22), followed by the North (2.09) and then the South (1.00). To summarise, the extracted rules suggest that the thresholds for London are considerably higher than other parts of England.

Observation 8. *C5 in London remained significantly higher than other regions.*

Focusing on the consequent of tier at least 4, the rules with a single criterion in its antecedent were found only for C4 (see the second last column of Table 7). Here, the rules for the North and the South are showing almost identical threshold values (i.e. 22.67 and 22.63).

So far, we have discussed the rules involving only a single criterion in its antecedent. There also exist other rules involving multiple criteria in their antecedents. For example, the consequent of tier at least 3 shows that there are rules extracted with the criteria C4 and C5 together. The threshold values for these two criteria are 9.09 & 2.91 for the North, and 13.64 & 3.3 for the South, respectively. Please note that as these thresholds are paired together, the values cannot be compared individually. In this case, we can see that the pair of values for the North are both considerably smaller than the pair of values extracted for the South.

The downward rules are summarised in Table 8, and are visualised in Figure 8 for inspection. Looking at the first row showing thresholds for C1, there are five instances where South sans London has higher threshold values than the North, while there is one instance where North has higher threshold value. On

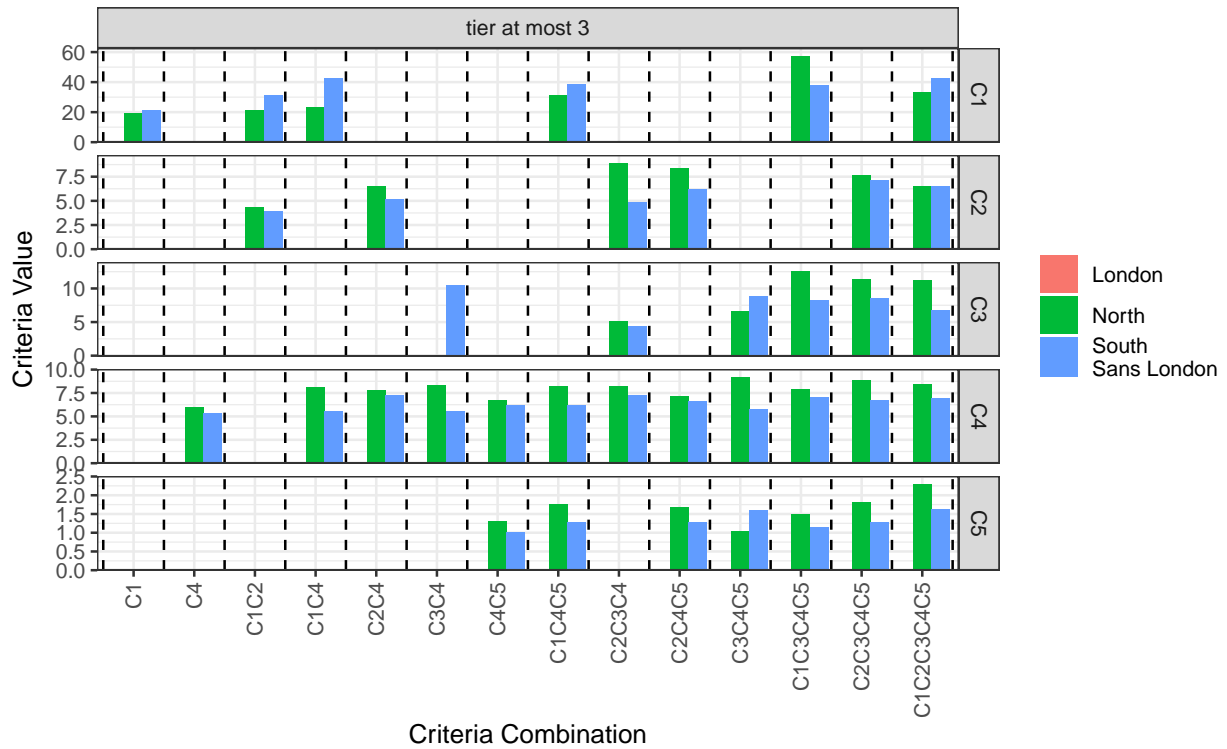


Figure 8: Downward Rules Plot

the contrary, looking at the second row, showing thresholds for C2, the North has higher threshold values in all cases. There are no rules extracted for London, however there are several rules extracted for North and South sans London, where, generally there is little to separate the threshold values for these two geographic regions.

Overall, given that the two geographical areas of the North and the South sans London are quite similar, and that London has more relaxed threshold values, it can be argued that although there is a North-South divide in terms of tiers allocation, it is one that is driven by London and not the other regions of the South. This results in London appearing to be given preferential treatment, with respect to the set of Criteria under consideration at least.

With London representing such a large percentage of the economic activity in England, this suggests that there may be additional implicit input into tier allocation reflecting economic considerations as additional criteria,

Observation 9. *There is a north south divide, in terms of tiers allocation, which is driven mostly by London.*

6. Conclusion

In this paper, we investigated the use of fairness in allocating tiered levels of restrictions, with the help of dominance-based rough sets analysis. We identify patterns in the data pertaining to the UK government criteria set for these tiered allocations which were translated into "if-then" type rules. We demonstrate a novel way of investigating fairness by extracting separate sets of rules for separate segments, and then comparing these rule sets to investigate fairness discrepancies in these segments. We found differences in the rules extracted from the North and the South of England, driven in large part by the region of London.

We found that there is a high level of correlation between C1 and C2, from which one could question how useful was it to have both in deciding the tiered restrictions. Intuitively, it makes sense to consider separate criteria to capture age disparities, given the consensus that age has a direct impact upon the severity of illness. However, our analysis suggests that the use of C2 as a separate criterion was statistically redundant and therefore not very useful. Our analysis also suggests that C3 was also not very useful, in a sense that it had almost no explanatory power in allocating tiers.

We argue that, despite being lauded as transparent, the proposed systematic approach was still not transparent enough. The UK government did collect the data on the set of criteria to assign tiers, but the use of this (and any other) information was obfuscated. As discussed earlier, the UK government gave consideration to the local context and exceptional circumstances such as a local but contained outbreak. This information was however not quantifiable and therefore cannot be included in a numerical model, however, it might be a possible area of future work involving the use of text mining techniques.

There is no information made public regarding the relative importance assigned to each of these criteria. As the problem involved multiple criteria, the focus of released information remained on public health which ignored the use of other important information related to, for example, economy, society and technology. Essentially, the allocation of tiers is a problem involving conflicting objectives where minimising the risk conflicts with the maximising of economic prosperity. Therefore, this is an area of research involving trade-off analysis and other multi-criteria decision making techniques.

One of the implications of a perceived lack of transparency of such a system, combined with an apparent disparity in treatment for different segments, could lead to a breakdown of trust. This in turn could lead to non-conformity to the rules, thus defeating the purpose of the whole tiered-allocation system.

References

- Alicke, M. D., Zell, E., 2008. Social comparison and envy. In: Smith, R. H. (Ed.), *Envy: Theory and research*. Oxford University Press, pp. 73–93.
- Alkaabneh, F., Diabat, A., Gao, H. O., 4 2021. A unified framework for efficient, effective, and fair resource allocation by food banks using an Approximate Dynamic Programming approach. *Omega* 100, 102300.
- Bassiouni, M. M., Chakraborty, R. K., Hussain, O. K., Rahman, H. F., 1 2023. Advanced deep learning approaches to predict supply chain risks under COVID-19 restrictions. *Expert Systems with Applications* 211, 118604.
- Bhapkar, H. R., Mahalle, P. N., Shinde, G. R., Mahmud, M., 2021. Rough Sets in COVID-19 to Predict Symptomatic Cases. *Lecture Notes on Data Engineering and Communications Technologies* 60, 57–68.
URL <https://link.springer.com/chapter/10.1007/978-981-15-9682-7>
- Błaszczczyński, J., de Almeida Filho, A. T., Matuszyk, A., Szeląg, M., Słowiński, R., 1 2021. Auto loan fraud detection using dominance-based rough set approach versus machine learning methods. *Expert Systems with Applications* 163, 113740.
- Błaszczczyński, J., Słowiński, R., Szeląg, M., 3 2011. Sequential covering rule induction algorithm for variable consistency rough set approaches. *Information Sciences Informatics and Computer Science, Intelligent Systems, Applications: An International Journal* 181 (5), 987–1002.
URL <https://dl.acm.org/doi/abs/10.1016/j.ins.2010.10.030>
- Busch, G., Spiller, A., 2016. Farmer share and fair distribution in food chains from a consumer’s perspective. *Journal of Economic Psychology* 55, 149–158.
- Carthey, J., Walker, S., Deelchand, V., Vincent, C., Griffiths, W. H., 2011. Breaking the rules: understanding non-compliance with policies and guidelines. *BMJ* 343.
URL <https://www.bmj.com/content/343/bmj.d5283>
- Department of Health and Social Care, GOV.UK, 2020. Tiering decisions to be made through national command structure.
URL <https://www.gov.uk/government/news/tiering-decisions-to-be-made-through-national-command-structure>
- DHSC, Hancock, M., 2020. Written statement to Parliament - Returning to a regional tiered approach.
URL <https://www.gov.uk/government/speeches/returning-to-a-regional-tiered-approach>
- Figueiredo, C. J. J. d., Mota, C. M. d. M., Rosa, A. G. F., Souza, A. P. G. d., Lima, S. M. d. S., 2022. Vulnerability to COVID-19 in Pernambuco, Brazil: A geospatial evaluation supported by multiple-criteria decision aid methodology. *Geospatial Health* 17 (s1).
URL <https://www.geospatialhealth.net/index.php/gh/article/view/1000>
- Gaile, G. L., 1977. Effiquity: A Comparison of a Measure of Efficiency with an Entropic Measure of the Equality of Discrete Spatial Distributions. *Economic Geography* 53 (3), 265–282.
- GOV.UK, 2020a. Analysis of the health, economic and social effects of COVID-19 and the approach to tiering Executive Summary. Tech. rep., HM Government, UK.
- GOV.UK, 2020b. The Health Protection (Coronavirus, Restrictions) (All Tiers) (England) Regulations 2020. Tech. rep., Public Health England.
- Greco, S., Matarazzo, B., Slowinski, R., 2001. Rough sets theory for multicriteria decision analysis. *European Journal of Operation Research* 129 (1), 1–47.
- Greco, S., Matarazzo, B., Slowinski, R., Stefanowski, J., 10 2000. An Algorithm for Induction of Decision Rules Consistent with

- the Dominance Principle. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2005, 304–313.
- Katris, C., 3 2021. A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece. *Expert Systems with Applications* 166, 114077.
- Khan, M., Mehran, M. T., Haq, Z. U., Ullah, Z., Naqvi, S. R., Ihsan, M., Abbass, H., 12 2021. Applications of artificial intelligence in COVID-19 pandemic: A comprehensive review. *Expert Systems with Applications* 185, 115695.
- Liou, J. J., Tang, C. H., Yeh, W. C., Tsai, C. Y., 10 2011. A decision rules approach for improvement of airport service quality. *Expert Systems with Applications* 38 (11), 13723–13730.
- Moss, R., 2020. Why People Aren't Going To Follow The Three-Tier Lockdown Rules. Tech. rep., HuffPost UK.
URL <https://bit.ly/3qWJ8xr>
- Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C., Vasilakis, C., 2021. Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. *European Journal of Operational Research* 290 (1), 99–115.
- Norheim, O. F., 12 2016. Ethical priority setting for universal health coverage: challenges in deciding upon fair distribution of health services. *BMC Medicine* 14 (1), 75.
URL <http://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-016-0624-4>
- ONS, 2016. England. Detailed information on the administrative structure within England.
URL www.ons.gov.uk/methodology/geography/ukgeographies/administrativegeography/england
- Shankar, R., Shah, A., 2003. Bridging the Economic Divide Within Countries: A Scorecard on the Performance of Regional Policies in Reducing Regional Income Disparities. *World Development* 31 (8), 1421–1441.
- Smith, M., 2018. What regions make up the North and South of England? | YouGov.
URL <https://yougov.co.uk/topics/politics/articles-reports/2018/01/04/what-regions-make-north-and-south-england>
- Tavana, M., Govindan, K., Nasr, A. K., Heidary, M. S., Mina, H., 2021. A mathematical programming approach for equitable COVID-19 vaccine distribution in developing countries. *Annals of Operations Research*.
- The Economist, 2012. The great divide.
URL <https://www.economist.com/britain/2012/09/15/the-great-divide>
- Webb, J., Johns, M., Roscoe, E., Giovannini, A., Qureshi, A., Baldini, R., 2022. STATE OF THE NORTH 2021/22 POWERING NORTHERN EXCELLENCE The progressive policy think tank. Tech. rep.
URL www.ippr.org/north
- Williams, S. N., Armitage, C. J., of Health Psychology, P., Dienes, K., in Clinical, L., Psychology, H., 11 2020. Public perceptions of non-adherence to COVID-19 measures by self and others in the United Kingdom. medRxiv, 2020.11.17.20233486.
URL <https://www.medrxiv.org/content/10.1101/2020.11.17.20233486v1><https://www.medrxiv.org/content/10.1101/2020.11.17.20233486v1.abstract>