



**Untersuchung des Zusammenhangs zwischen der  
Dunkelziffer und der Teststrategie in der  
Ausbreitung von COVID-19 in Europa**

Ludwig-Maximilians-Universität München  
Institut für Statistik

**Abschlussarbeit**

zur Erlangung des akademischen Grades  
Bachelor of Science

**Gjeneta Ramadani**

1.Juli 2021

**Betreuer:** Dr. Sabine Hoffmann

## **Eidesstattliche Erklärung**

Hiermit versichere ich, die vorliegende Abschlussarbeit selbstständig und nur unter Verwendung der von mir angegebenen Quellen und Hilfsmittel verfasst zu haben. Sowohl inhaltlich als auch wörtlich entnommene Inhalte wurden als solche kenntlich gemacht. Die Arbeit hat in dieser oder vergleichbarer Form noch keinem anderem Prüfungsgremium vorgelegen.

Datum: \_\_\_\_\_ Unterschrift: \_\_\_\_\_

## Abstract

Die vorliegende Arbeit zeigt unterschiedliche Ansätze zur Modellierung des Zusammenhangs zwischen der Dunkelziffer, der Testpositivrate und der Testgröße in der Ausbreitung von COVID-19 in den fünf europäischen Ländern Deutschland, Frankreich, Spanien, Italien und Dänemark. Dabei liegt der Fokus hauptsächlich auf Deutschland. Zunächst werden die Methoden hinter der Schätzung der Infektionszahlen nur auf Basis von gemeldeten Todesfällen aufgezeigt. Es folgt eine Einführung in die Theorie der Zeitreihenanalyse, auf deren Grundlage verschiedene statische und dynamische Modelle zur Analyse des Zusammenhangs von Dunkelziffer, Testgröße und Positivrate aufgestellt werden.

Es zeigt sich, dass die dynamischen Modelle nicht zur Analyse dieses Zusammenhangs geeignet sind. Die statischen Modelle sind besser geeignet, verletzen jedoch zwei zentrale Annahmen der Zeitreihenregressionsanalyse. Die Autokorrelation der Fehler kann mithilfe alternativer Neuschätzungen der Standardabweichungen der Koeffizienten adressiert werden. Die Behebung der Autokorrelation der statischen Modelle mithilfe von Differenzenbildung stellt sich aufgrund der Eliminierung langfristiger Zusammenhänge als ungeeignet heraus.

Abschließend werden für den Einfluss der Positivrate Dummy-Variablen zur Berücksichtigung von Strukturbrüchen in der Entwicklung der Pandemie benutzt. Die Ergebnisse der Modellgleichungen für alle fünf betrachteten Länder besagen, dass mithilfe der Positivrate die Zahl der Infektionen gut modelliert werden kann, und dass auch der relative Anteil der Tests an den Infektionen einen signifikanten negativen Einfluss auf den Anteil unentdeckter Fälle an den Infektionen ausübt.

# Inhaltsverzeichnis

Eidesstattliche Erklärung . . . . .	2
Abstract . . . . .	3
<b>1 Einleitung</b>	<b>9</b>
1.1 Motivation . . . . .	9
1.2 Datengrundlage . . . . .	11
1.2.1 Datenquellen . . . . .	11
1.2.2 Datenaufbereitung . . . . .	12
<b>2 Methoden zur Bestimmung der Infektionszahlen und zur allgemeinen Zeitreihenanalyse</b>	<b>14</b>
2.1 Berechnung der Infektionszahlen . . . . .	14
2.1.1 Berechnung der infection fatality rates . . . . .	14
2.1.2 Gammaverteilung . . . . .	15
2.1.3 Abschließende Berechnung der Infektionszahlen . . . . .	17
2.2 Reporting Delay . . . . .	20
2.3 Theorie der Zeitreihenanalyse . . . . .	22
2.3.1 Einführung: Das einfache lineare Modell . . . . .	22
2.3.2 Zeitreihenanalyse: Statische vs. Dynamische Modelle . . . . .	23
2.4 Evaluationen . . . . .	26
2.4.1 Teil I: Autokorrelation der Fehler . . . . .	26
2.4.2 Teil II: Stationarität der Variablen . . . . .	29
2.5 Aufstellung der Modellgleichungen . . . . .	30
2.5.1 Problematiken des Zusammenhangs zwischen Infektionen, Dun- kelziffer, Positivrate und Testgröße . . . . .	30
2.5.2 Modellgleichungen . . . . .	32
2.6 Exkurs: Spezifität und Sensitivität . . . . .	35
2.6.1 Einführung und Definitionen . . . . .	35
2.6.2 Methode zur Adjustierung der Fallzahlen . . . . .	35
<b>3 Ergebnisse</b>	<b>38</b>
3.1 Deskriptive Analyse von Testgröße, gemeldeten Fällen und Positivrate .	38
3.2 Berechnung der Infektionszahlen . . . . .	41
3.3 Reporting Delay . . . . .	45

3.4	Statische Modelle . . . . .	47
3.4.1	Ein naives Modell . . . . .	47
3.4.2	Lösungsansätze für den Einfluss der Testgröße . . . . .	48
3.4.3	Lösungsansätze für den Einfluss der Positivrate . . . . .	56
3.5	Evaluationen der statischen Modelle . . . . .	62
3.5.1	Teil I: Autokorrelation der Fehler . . . . .	62
3.5.2	Teil II: Stationarität der Variablen . . . . .	66
3.6	Dynamische Modelle . . . . .	73
3.6.1	Dynamisches Testgrößenmodell . . . . .	73
3.6.2	Dynamisches Positivrate-Infektionen-Modell . . . . .	75
3.7	Betrachtung weiterer europäischer Länder . . . . .	77
3.7.1	Frankreich . . . . .	78
3.7.2	Spanien . . . . .	82
3.7.3	Italien . . . . .	85
3.7.4	Dänemark . . . . .	88
<b>4</b>	<b>Diskussion</b>	<b>91</b>
4.1	Zusammenfassende Bewertung . . . . .	91
4.1.1	Zusammenfassung der Modellergebnisse . . . . .	91
4.1.2	Zusammenfassung der methodischen Hintergründe . . . . .	92
4.2	Ausblick . . . . .	93
4.2.1	Spezifität und Sensitivität . . . . .	94

# Abbildungsverzeichnis

2.1	Flussdiagramm zu den Zusammenhängen zwischen Infektionen, Dunkelziffer, Positivrate und Testgröße . . . . .	30
3.1	Positivrate in Deutschland . . . . .	38
3.2	Anzahl Tests in Deutschland . . . . .	39
3.3	Tests und gemeldete Fälle in Deutschland . . . . .	40
3.4	Todesfälle in Deutschland . . . . .	41
3.5	Modellierte Infektionen in Deutschland . . . . .	42
3.6	Infektionen vs. gemeldete Fälle in Deutschland . . . . .	43
3.7	Schätzung der Verteilung der Zeit zwischen Symptombeginn und Meldung . . . . .	45
3.8	Modellierte Infektionen mit reporting delay der Infektionsfälle in Deutschland . . . . .	46
3.9	Modellierte Infektionen mit reporting delay der Infektionsfälle und der Todesfälle in Deutschland . . . . .	46
3.10	Infektionen vs. Anzahl Tests in Deutschland . . . . .	48
3.11	Infektionen vs. Dunkelziffer in Deutschland . . . . .	49
3.12	Anzahl Tests vs. Dunkelziffer in Deutschland . . . . .	49
3.13	Anteil Dunkelziffer vs. Infektionen in Deutschland . . . . .	50
3.14	Anteil Tests und Anteil Dunkelziffer in Deutschland . . . . .	51
3.15	Modellschätzung durch Testgröße in Deutschland . . . . .	53
3.16	Streudiagramme der alternativen Definition der Dunkelziffer . . . . .	55
3.17	Infektionen vs. Positivrate in Deutschland . . . . .	57
3.18	Modellschätzung der Dunkelziffer durch Positivrate in Deutschland . . . . .	59
3.19	Modellschätzung der Infektionen durch Positivrate in Deutschland . . . . .	59
3.20	Streudiagramme der alternativen Definitionen der Positivrate und der Infektionen . . . . .	61
3.21	Residualanalyse der einfachen statischen Modelle . . . . .	64
3.22	Residuenplots der statischen Modelle . . . . .	65
3.23	Schätzungen der Differenzenmodelle . . . . .	67
3.24	Bildung erster Differenzen . . . . .	68
3.25	Modellschätzung durch Positivrate inkl. Strukturbrüche in Deutschland . . . . .	72

---

3.26	Modellschätzung durch Positivrate inkl. Strukturbrüche und Interaktionen in Deutschland . . . . .	72
3.27	Residuenanalyse des dynamischen Testgrößenmodells . . . . .	74
3.28	Residuenanalyse des dynamischen Positivrate-Infektionen-Modells . . .	76
3.29	Frankreich: Verlauf der Pandemie . . . . .	78
3.30	Modellschätzung der Infektionen durch Positivrate in Frankreich . . . .	80
3.31	Modellschätzung durch Testgröße in Frankreich . . . . .	81
3.32	Spanien: Verlauf der Pandemie . . . . .	82
3.33	Modellschätzung der Infektionen durch Positivrate in Spanien . . . . .	83
3.34	Modellschätzung durch Testgröße in Spanien . . . . .	84
3.35	Italien: Verlauf der Pandemie . . . . .	85
3.36	Modellschätzung der Infektionen durch Positivrate in Italien . . . . .	86
3.37	Modellschätzung durch Testgröße in Italien . . . . .	87
3.38	Dänemark: Verlauf der Pandemie . . . . .	88
3.39	Modellschätzung der Infektionen durch Positivrate in Dänemark . . . .	89
3.40	Modellschätzung durch Testgröße in Dänemark . . . . .	90
4.1	Einflüsse von Sensitivität und Spezifität in den Sommermonaten in Deutschland . . . . .	95
4.2	Einflüsse von Sensitivität und Spezifität während der zweiten Welle in Deutschland . . . . .	95

# Tabellenverzeichnis

3.1	Ergebnisse der <i>ifr</i> -Berechnung . . . . .	41
3.2	Ergebnisse des naiven Modells . . . . .	47
3.3	Ergebnisse des Testgrößenmodells . . . . .	52
3.4	Ergebnisse des Testgrößenmodells mit alternativer Definition der Dunkelziffer . . . . .	54
3.5	Ergebnisse des Positivrate-Dunkelziffer-Modells . . . . .	56
3.6	Ergebnisse des Positivrate-Infektionen-Modells . . . . .	57
3.7	Newey-West-Schätzung, Testgrößenmodell . . . . .	63
3.8	gewöhnliche Schätzung, Testgrößenmodell . . . . .	63
3.9	Newey-West-Schätzung, Positivrate-Infektionen-Modell . . . . .	63
3.10	gewöhnliche Schätzung, Positivrate-Infektionen-Modell . . . . .	63
3.11	Ergebnisse des Differenzen-Testgrößenmodells . . . . .	66
3.12	Ergebnisse des Differenzen-Positivrate-Infektionen-Modells . . . . .	66
3.13	Ergebnisse des Positivrate-Infektionen-Modells mit Strukturbrüchen . . . . .	70
3.14	Ergebnisse des Positivrate-Infektionen-Modells mit Strukturbrüchen . . . . .	71
3.15	Ergebnisse des dynamischen Testgrößenmodells . . . . .	73
3.16	Ergebnisse des dynamischen Positivrate-Infektionen-Modells . . . . .	75
3.17	Frankreich: Ergebnisse des Positivrate-Infektionen-Modells . . . . .	79
3.18	Frankreich: Ergebnisse des Testgrößenmodells . . . . .	81
3.19	Spanien: Ergebnisse des Positivrate-Infektionen-Modells . . . . .	83
3.20	Spanien: Ergebnisse des Testgrößenmodells . . . . .	84
3.21	Italien: Ergebnisse des Positivrate-Infektionen-Modells . . . . .	86
3.22	Italien: Ergebnisse des Testgrößenmodells . . . . .	87
3.23	Dänemark: Ergebnisse des Positivrate-Infektionen-Modells . . . . .	89
3.24	Dänemark: Ergebnisse des Testgrößenmodells . . . . .	90

# 1 Einleitung

## 1.1 Motivation

In der Bekämpfung der Corona-Pandemie nehmen die gemeldeten Fallzahlen der mit dem neuartigen Coronavirus SARS-CoV-2 infizierten Personen eine Schlüsselrolle ein. Nicht nur dienen sie der Überwachung des aktuellen Infektionsgeschehens, sondern sie bilden auch die Grundlage politischer Entscheidungen und Maßnahmen zur Eindämmung der Pandemie. In Deutschland gelten bundesweit Ausgangssperren ab einer 7-Tage-Inzidenz von 100 (100 gemeldete Fälle pro 100 000 Einwohner innerhalb von sieben Tagen), Kontaktbeschränkungen und Schließungen ganzer Wirtschaftsbereiche wie der Gastronomie und der Kulturbranche (Stand: Mai 2021). Lockerungen dieser politischen Maßnahmen werden an fallende Fallzahlen gekoppelt, die wiederum durch gemeldete Fallzahlen gespiegelt werden sollen (vgl. [Baden-Württemberg, 2021]).

Die Interpretation der gemeldeten Fallzahlen sollte jedoch mit Vorsicht vorgenommen werden. Es gibt viele Unsicherheiten und Herausforderungen bei der alleinigen Interpretation dieser Zahlen. Beispielsweise können diese nur einen Teil des Infektionsgeschehens widerspiegeln, da immer nur ein Teil der Bevölkerung sich überhaupt testen lässt. Somit sind die gesammelten Testergebnisse als eine Art Stichprobe aus der Gesamtbevölkerung zu verstehen, die womöglich nicht repräsentativ für das wahre Infektionsgeschehen ist. Es muss von unentdeckten Fällen ausgegangen werden, der sogenannten Dunkelziffer. Zudem sind Tests auch immer mit einem Fehler behaftet, falsch-positive sowie falsch-negative Ergebnisse können womöglich auch das Bild der pandemischen Entwicklung verzerren.

Diese Arbeit analysiert diese Problematiken. Sie untersucht den Zusammenhang zwischen einer variierenden Testgröße und einer sich fortlaufend ändernden Positivrate der durchgeführten Test auf der einen Seite, und der Dunkelziffer und darüber hinaus auch den Infektionszahlen auf der anderen Seite. Es werden Wege aufgezeigt, wie Infektionszahlen nur mithilfe von gemeldeten Todesfällen geschätzt werden können, um das wahre Infektionsgeschehen offen zu legen. Außerdem wird eine Möglichkeit aufgezeigt, wie mit Verzerrungen der gemeldeten Infektionsfälle aufgrund eines Meldeverzugs umgegangen werden kann. Zusätzlich wird diskutiert, wie sich falsch negative und falsch positive Tests auf die gemeldeten Infektionszahlen auswirken können.

All diese Grenzen der bloßen Betrachtung von gemeldeten Infektionszahlen, die Verzögerung der Meldung von Infektionsfällen, die Fehleranfälligkeit von Tests und eine ständig variierende Zahl nicht entdeckter Infektionen, werden in dieser Arbeit analysiert. Der Hauptfokus liegt jedoch ganz klar in der Schätzung der wahren Infektionszahlen bzw. der Dunkelziffer anhand zweier Eigenschaften des Testprozesses, nämlich der Positivrate und der Testgröße.

## 1.2 Datengrundlage

### 1.2.1 Datenquellen

Als Datengrundlage für die im weiteren Verlauf dieser Arbeit angewandten Analysen dienen publizierte Daten des *Europäischen Zentrums für die Prävention und die Kontrolle von Krankheiten*, kurz *ECDC* (*European Centre for Disease Prevention and Control*), und publizierte Daten der *John-Hopkins-Universität*, kurz *JHU*.

Das *ECDC* ist eine europäische Agentur mit dem Ziel, die Abwehr infektiöser Krankheiten in Europa zu stärken [ECDC, 2021a]. Zu ihren Aufgaben zählen unter Anderem die Ermittlung, Bewertung und Kommunikation von durch infektiöse Krankheiten verursachten Bedrohungen für die menschliche Gesundheit und die Information der europäischen Mitgliedstaaten und der europäischen Kommission darüber [ECDC, 2021f]. Es wurden zwei von dem *ECDC* bereitgestellte Datensätze für diese Arbeit verwendet, einmal der Datensatz zu täglich gemeldeten Fällen von COVID-19 (coronavirus disease 2019) und zu COVID-19-Todesfällen weltweit [ECDC, 2021b].

Diese Daten werden von dem *ECDC* seit Beginn der Pandemie gesammelt, genauer dem 31. Dezember 2019, basierend auf den Informationen von Gesundheitsbehörden weltweit. Dieser Prozess endete am 14. Dezember 2020, wurde allerdings ab dem 11. März 2021 wieder fortgesetzt. In dieser Arbeit wird lediglich der archivierte Datensatz mit Daten bis zum 14. Dezember verwendet. Für die Daten aus Ländern der EU bzw. des EEA (EU-Länder plus Liechtenstein, Island und Norwegen) überprüft das *ECDC* täglich zwischen 8 und 10 Uhr mitteleuropäischer Zeit bis zu 100 relevante Datenquellen. Diese werden mit Informationen von der Weltgesundheitsorganisation (WHO), dem Frühwarn- und Reaktionssystem (EWRS) und dem europäischen Überwachungssystem (TESSy) vervollständigt [ECDC, 2021c]. Tägliche Updates und Visualisierungen dieses Datensatzes können auf der ECDC-Website eingesehen werden.

Der zweite Datensatz des *ECDC*, der in dieser Arbeit verwendet wird, enthält wöchentliche Informationen über die Anzahl durchgeführter Tests, die Testpositivrate und die wöchentlich gemeldeten Fallzahlen über Länder der EEA [ECDC, 2021d]. Auch diese Daten werden von nationalen Gesundheitsbehörden und Gesundheitsministerien gesammelt und mit den Daten der WHO überprüft. Die Hauptquelle für die Testgröße stellt dabei das TESSy dar, wobei bei fehlenden Informationen weitere öffentliche Online-Quellen genutzt werden [ECDC, 2021e].

Da das *ECDC* zwischen dem 14. Dezember 2020 und dem 11. März 2021 keine täglichen Daten zu Todesfällen bereitstellt, werden Informationen über Todesfälle nach dem 14. Dezember von einer weiteren Datenquelle, der *JHU* herangezogen [GitHub, 2021b]. Diese aggregiert Daten aus zahlreichen Quellen, darunter der Weltgesundheitsorganisation, nationalen Gesundheitsbehörden, dem COVID Tracking Project uvm. Eine ausführliche Liste aller Quellen findet man unter <https://github.com/CSSEGISandData/COVID-19>. Die *JHU* aktualisiert stündlich ihr Dashboard mit weltweiten Daten über die Ausbreitung von COVID-19 [JHU.edu, 2021].

Zusätzlich werden später bei der Infektionszahlberechnung, genauer bei der Berechnung der *infection fatality rates*, Daten zur altersspezifischen Bevölkerungszusammensetzung und zu altersspezifischen sozialen Interaktionen der betrachteten Nationen benötigt. Dazu werden zum Einen Daten der Vereinten Nationen [UnitedNations, 2019] gebraucht. Dieser Datensatz enthält solche Informationen über alle UN-Mitgliedstaaten. Er ist in Abschnitten von Altersgruppen aufgeteilt, wobei jede Gruppe eine Altersspanne von fünf Jahren umfasst. Er zeigt für jedes Land, wie hoch die Anzahl der 0-4-Jährigen, 5-9-Jährigen usw. bis zu der Gruppe der 95-99-Jährigen ist. Die letzte Gruppe beschreibt die Altersgruppe der über 100-Jährigen.

Zum Anderen erhält man über das R-Paket *socialmixr* [GitHub, 2021a] Informationen über die Kontaktmuster acht verschiedener europäischer Länder, darunter auch Deutschland, Italien und die Niederlande. Die Daten stammen aus einer Studie aus dem Jahre 2008 mit 7290 Teilnehmern, die Charakteristika ihrer Kontakte, wie Häufigkeiten, Dauer und Altersgruppe, aufgezeichnet haben [Mossong et al., 2008]. Diese Informationen über die Kontaktanteile zwischen Individuen verschiedener Altersgruppen werden bei der Berechnung der sogenannten *infection fatality rates* gebraucht. Darauf wird in Kapitel (2.1.1) näher eingegangen.

## 1.2.2 Datenaufbereitung

Der in dieser Arbeit verwendete Datensatz über tägliche Todeszahlen des *ECDC* enthält Daten bis zum 14. Dezember 2020, der der *JHU* bis zum 26. April 2021. Der Datensatz über die Teststrategien, also über die Positivrate und die Testgröße, des *ECDC* enthält wöchentliche Daten ab dem Februar 2020 für die Länder Frankreich, Dänemark, Italien und Deutschland, und ab Ende April 2020 auch für Spanien. Die Daten für die Teststrategien laufen für alle fünf Länder bis Ende Februar 2021. Wir

betrachten also insgesamt einen Zeitraum von ungefähr einem Jahr, nämlich von Februar 2020 bis Februar 2021. Zusätzlich betrachten wir nur die fünf uns interessierenden Länder Deutschland, Frankreich, Spanien, Italien und Dänemark.

Bei der Zusammenführung aller drei Datensätze wird wie folgt vorgegangen. Die beiden Datensätze über die täglichen Todesfälle werden zusammengeführt, wobei zu beachten ist, dass für Spanien der ECDC-Datensatz keine Daten für den 14. Dezember 2020 enthält, für die restlichen vier Länder hingegen schon. Dieser fehlende Tag wird mit den Daten aus dem JHU-Datensatz vervollständigt. Der Datensatz der *ECDC* enthält drei negative Einträge über die Todesfälle, zwei Tage Spaniens und einen Tag Italiens. Diese drei Einträge werden bereinigt, indem für jeden dieser Tage das arithmetische Mittel aus dem vorangegangenen und dem nachfolgenden Tag berechnet wird. Erst nach der Berechnung der Infektionszahlen auf Grundlage dieses zusammengeführten Datensatzes, werden die berechneten Infektionszahlen auf wöchentlicher Basis aggregiert, damit sie zu dem Teststrategien-Datensatz passen. Auch wenn sowohl der JHU-Datensatz als auch der ECDC-Datensatz Informationen zu den gemeldeten Fällen beinhalten, werden die Daten über diese gemeldeten Fälle aus dem Teststrategien-Datensatz des *ECDC* genommen, da diese genau zu der Testgröße und der Testpositivrate dieses Datensatzes passen.

Für die Kontaktmuster der fünf in dieser Arbeit betrachteten Länder werden aus dem *socialmixr*-Paket die bereitgestellten Informationen für Deutschland und Italien entnommen. Da Kontaktdaten für die Länder Frankreich, Spanien und Dänemark fehlen, werden sie durch repräsentative Kontaktmuster ersetzt. Analog wie in Flaxman et al. [2020], wird dafür das Kontaktmuster der Niederlande verwendet.

# 2 Methoden zur Bestimmung der Infektionszahlen und zur allgemeinen Zeitreihenanalyse

## 2.1 Berechnung der Infektionszahlen

Eine besondere Herausforderung dieser Arbeit stellt die angemessene Schätzung der tatsächlichen Infektionszahlen dar. Dafür werden lediglich Daten über gemeldete Todesfälle nach einer Infektion mit dem Coronavirus benutzt. Gemeldete Fallzahlen werden nicht betrachtet. Ein Vorteil dieser Methode besteht darin, dass wir uns letztendlich für den Unterschied zwischen Infektionszahlen und gemeldeten Fällen interessieren, und die auf diese Weise geschätzten Infektionen komplett unberührt von den gemeldeten Fällen sind. Zudem kann man davon ausgehen, dass die Zahl gemeldeter Todesfälle zuverlässiger als die der Fallzahlen ist, da sie strenger der Kontrolle von Gesundheitsbehörden und Krankenhäusern unterliegen.

Die Schätzung basiert auf zwei festen Schätzungen epidemiologischer Parameter, der *infection fatality rate* und der *Verteilung der Dauer zwischen dem Infektionszeitpunkt und dem Todeszeitpunkt*. Die *infection fatality rate* gibt an, wie hoch die Wahrscheinlichkeit ist, an einer Infektion zu sterben, unter der Bedingung, dass eine Infektion vorliegt. Im Folgenden wird näher auf die Bestimmung beider Parameter eingegangen.

### 2.1.1 Berechnung der infection fatality rates

Die altersspezifischen *infection fatality rates*  $ifr_a$  werden aus Verity et al. [2020] entnommen. Diese nutzen hierfür Todesfälle bis zum 8. Februar 2020 aus China. Dieser Schätzung liegt eine gleichverteilte Verteilungsannahme der *Befallsraten* in allen Altersgruppen zugrunde [Verity et al., 2020]. Dabei wird eine perfekte Erfassung der Fallzahlen in der Altersgruppe von 50 - 59 Jahren angenommen. Genau wie in Flaxman et al. [2020] werden die unangepassten  $ifr'$  zunächst angepasst an bessere Schätzungen der *Befallsraten*  $c$ . Sei  $c_a$  die Anzahl der Infektionen in der Altersgruppe  $a$ ,  $N_a$  die Größe dieser Altersgruppe, so ergibt sich die *Befallsrate*  $BR$  durch  $BR_a = \frac{c_a}{N_a}$ . Die

adjustierten  $ifr$  berechnen sich folgendermaßen:

$$ifr_a = \frac{BR_{50-90}}{BR_a} \cdot ifr'_a \quad (2.1)$$

Diese Anpassung ergibt sich aus folgenden Umformungen, wobei  $d_a$  die Anzahl der Todesfälle in der Altersgruppe  $a$  widerspiegelt:

$$ifr'_a = \frac{d_a}{\frac{c_{50-59}}{N_{50-59}} \cdot N_a}$$

Der Nenner zeigt dabei die Anpassung der  $ifr'$  an die  $BR_{50-59}$

$$\begin{aligned} &= \frac{d_a}{\frac{c_{50-59}}{N_{50-59}} \cdot \frac{N_a}{c_a} \cdot c_a} \\ &= \frac{BR_a}{BR_{50-59}} \cdot \frac{d_a}{c_a} \end{aligned}$$

woraus die Gleichung aus (2.1) folgt.

Für die Schätzung der länderspezifischen  $AR_a$  aller betrachteten fünf Länder werden UN-Daten über die Bevölkerungszusammensetzung und Daten über Kontaktmuster aus dem *socialmixr-Paket* der verschiedenen Bevölkerungen herangezogen. Diese sagen aus, dass die *Befallsraten* tendenziell bei höheren Altersgruppen schrumpfen [Walker et al., 2020]. Für die Berechnung der länderspezifischen adjustierten  $ifr_a$  wird der in [Flaxman et al., 2020] benutzte Code genommen, welcher unter <https://github.com/ImperialCollegeLondon/covid19model/blob/master/compute-ifr-europe.r> öffentlich zugänglich ist.

### 2.1.2 Gammaverteilung

Die Schätzung der Infektionszahlen beruht auf Verteilungsannahmen über die Zeitdauer zwischen einer Infektion mit dem Coronavirus und dem Todeszeitpunkt. Diese Zeitdauer wird als aus zwei unabhängigen Verteilungen zusammengesetzt angenommen: die Inkubationszeit, also die Zeit zwischen einer Infektion und dem Auftreten von Symptomen, und die Zeit zwischen Auftreten von Symptomen und Todeszeitpunkt. Beide Verteilungen folgen einer Gamma-Verteilung. Die Inkubationszeit folgt dabei einer Gamma-Verteilung mit einem Erwartungswert von 5.1 Tagen und einem Variati-

onskoeffizienten von 0.86. Die Zeit zwischen dem Auftreten von Symptomen und dem Todeszeitpunkt ist auch gammaverteilt mit einem Erwartungswert von 17.8 Tagen und einem Variationskoeffizienten von 0.45 [Flaxman et al., 2020, S.4]. Bezeichne nun  $X_1$  die Verteilung der Inkubationszeit und  $X_2$  die Verteilung der Zeit zwischen Auftreten von Symptomen und Todeszeitpunkt, so ist die zeitliche Verteilung zwischen dem Zeitpunkt der Infektion und dem Zeitpunkt des Todes gegeben durch  $\Omega$ :

$$\begin{aligned} X_1 &\sim \text{Gamma}(5.1, 0.86) \\ X_2 &\sim \text{Gamma}(17.8, 0.45) \\ \Omega &= X_1 + X_2 \end{aligned} \tag{2.2}$$

In obiger Formel wurden die Gammaverteilungen mit ihren Erwartungswerten und Variationskoeffizienten parametrisiert. Üblicherweise erfolgt diese Parametrisierung mit den Parametern  $\alpha$  und  $\beta$ , auch als *shape* und *rate* bezeichnet. Es gelten folgende Beziehungen einer gammaverteilten Variable  $X$  mit  $X \sim \text{Gamma}(\alpha, \beta)$ :

$$E(X) = \frac{\alpha}{\beta} \tag{2.3}$$

$$\text{Var}(X) = \frac{\alpha}{\beta^2} \tag{2.4}$$

$$\begin{aligned} E(\Omega) &= E(X_1) + E(X_2) \\ &= 5.1 + 17.8 \\ &= 22.9 \end{aligned} \tag{2.5}$$

Unter der Unabhängigkeitsannahme von  $X_1$  und  $X_2$  gilt zudem:

$$\text{Var}(\Omega) = \text{Var}(X_1) + \text{Var}(X_2) \tag{2.6}$$

Mithilfe der bekannten Variationskoeffizienten  $\text{Var}K$  von  $X_1$  und  $X_2$  lässt sich auch die Varianz von  $\Omega$  bestimmen:

$$\begin{aligned} \text{Var}K(X_1) &= \frac{\sqrt{\text{Var}(X_1)}}{E(X_1)} \\ \Leftrightarrow \text{Var}(X_1) &= (\text{Var}K(X_1) \cdot E(X_1))^2 \\ &= (0.86 \cdot 5.1)^2 \\ &\approx 19.2370 \end{aligned} \tag{2.7}$$

Analog dazu gilt:

$$\begin{aligned} \text{Var}(X_2) &= (0.45 \cdot 17.8)^2 \\ &\approx 64.1601 \end{aligned}$$

Nach (2.6) gilt somit  $\text{Var}(\Omega) \approx 83.3971$ . Abschließend lassen sich mithilfe von (2.3) und (2.4) die Parameter *shape* und *rate* von  $\Omega$  berechnen und es gilt:

$$\Omega \sim \text{Gamma}(6.2881, 0.2746)$$

### 2.1.3 Abschließende Berechnung der Infektionszahlen

Die abschließende Berechnung der Infektionszahlen erfolgt in zwei Schritten. Aus den gegebenen Todeszahlen wird rückwirkend die Anzahl der Infektionen, die innerhalb von 60 Tagen zum Tod führen werden, berechnet. Das Zeitfenster von 60 Tagen wurde gewählt, da  $P(\Omega \leq 60) \approx 0.9986$ , also fast eins, und somit so gut wie alle Infektionen, die zum Tod führen werden, abgedeckt sind. Dann können mithilfe der bereits geschätzten *ifr* aller Länder, die tatsächlichen täglichen Infektionszahlen berechnet werden. Die erwartete Anzahl an Infektionen  $r$ , die zum Tod führen werden, sind für Tag  $t$  und Land  $m$  durch folgende Summe gegeben:

$$r_{t,m} = \sum_{i=1}^{60} d_{t+i,m} \cdot (P(\Omega \leq i + 0.5) - P(\Omega \leq i - 0.5)) \quad (2.8)$$

Die Zahlen 0.5 und  $-0.5$  in der Formel (2.8) wurden gewählt, um einen ganzen Tag abzudecken. Die Überlegung hinter der Zahl 0.5 sieht so aus, dass bei gleichförmiger Verteilung der Todeszeitpunkte innerhalb eines Tages davon ausgegangen werden kann, dass im Mittel der Todeszeitpunkt auf die Mitte des Tages fällt. Somit berechnet sich bspw. die Wahrscheinlichkeit, sich einen Tag ( $i = 1$ ) vor dem Todestag infiziert zu haben, als die Wahrscheinlichkeit, zwischen einem halben Tag ( $i - 0.5 = 0.5$ ) und eineinhalb Tagen ( $i + 0.5 = 1.5$ ) vor dem Todestag eine Infektion zu erleiden. Für die Berechnung von  $r_{0,m}$  werden die Todeszahlen des Folgetages ( $t = 1$ ) mit der Wahrscheinlichkeit, innerhalb eines Tages nach einer Infektion zu sterben, multipliziert. Die Todeszahlen des darauffolgenden Tages werden mit der Wahrscheinlichkeit, zwischen eineinhalb und zweieinhalb Tagen zu sterben, multipliziert und dazuaddiert. Es ergibt

sich die Formel:

$$\begin{aligned}
 r_{0,m} &= d_{1,m} \cdot (P(\Omega \leq 1.5) - P(\Omega \leq 0.5)) \\
 &+ d_{2,m} \cdot (P(\Omega \leq 2.5) - P(\Omega \leq 1.5)) \\
 &+ d_{3,m} \cdot (P(\Omega \leq 3.5) - P(\Omega \leq 2.5)) \\
 &+ \dots \\
 &+ d_{60,m} \cdot (P(\Omega \leq 60.5) - P(\Omega \leq 59.5)) \\
 &= \sum_{i=1}^{60} d_{i,m} \cdot (P(\Omega \leq i + 0.5) - P(\Omega \leq i - 0.5))
 \end{aligned}$$

Für die Berechnung der Infektionszahlen die zum Tod führen für Land  $m$  an Tag 1 gilt analog:

$$\begin{aligned}
 r_{1,m} &= d_{2,m} \cdot (P(\Omega \leq 1.5) - P(\Omega \leq 0.5)) \\
 &+ d_{3,m} \cdot (P(\Omega \leq 2.5) - P(\Omega \leq 1.5)) \\
 &+ \dots \\
 &+ d_{61,m} \cdot (P(\Omega \leq 60.5) - P(\Omega \leq 59.5)) \\
 &= \sum_{i=2}^{61} d_{i+1,m} \cdot (P(\Omega \leq i - 0.5) - P(\Omega \leq i - 1.5))
 \end{aligned}$$

oder auch:

$$= \sum_{i=1}^{60} d_{1+i,m} \cdot (P(\Omega \leq i + 0.5) - P(\Omega \leq i - 0.5))$$

und für  $t = 2$ :

$$\begin{aligned}
 r_{2,m} &= d_{3,m} \cdot (P(\Omega \leq 1.5) - P(\Omega \leq 0.5)) \\
 &+ d_{4,m} \cdot (P(\Omega \leq 2.5) - P(\Omega \leq 1.5)) \\
 &+ \dots \\
 &+ d_{62,m} \cdot (P(\Omega \leq 60.5) - P(\Omega \leq 59.5)) \\
 &= \sum_{i=3}^{62} d_{i,m} \cdot (P(\Omega \leq i - 1.5) - P(\Omega \leq i - 2.5))
 \end{aligned}$$

oder auch:

$$= \sum_{i=1}^{60} d_{2+i,m} \cdot (P(\Omega \leq i + 0.5) - P(\Omega \leq i - 0.5))$$

Es folgt also (2.8). Abschließend muss, um die Zahl der tatsächlichen Infektionen  $c$  für Zeitpunkt  $t$  und Land  $m$  zu bestimmen,  $r_{t,m}$  durch die  $ifr$  von Land  $m$ , also durch  $ifr_m$  geteilt werden. Somit gilt:

$$c_{t,m} = \frac{r_{t,m}}{ifr_m} \tag{2.9}$$

## 2.2 Reporting Delay

Bei der Analyse der Dunkelziffer ist es nicht akkurat, die geschätzten Infektionszahlen mit den zum jeweiligen Datum gemeldeten Fallzahlen direkt zu vergleichen. Der Meldeverzug, also die Verzögerung zwischen dem Auftreten einer Infektion und dem offiziellen Melden dieser, muss berücksichtigt werden. Dr. Sabine Hoffmann, Betreuerin dieser Arbeit, hat mit ihrem Modell, unter Berücksichtigung bayerischer Daten, durch das Beobachten dieser Zeitintervalle, Schätzungen für den *reporting delay* berechnet und für diese Arbeit zur Verfügung gestellt. Der *reporting delay* beschreibt dabei die Zeit zwischen dem Einsetzen der Symptome und dem Melden als positives Testergebnis. Folglich muss zusätzlich noch die Inkubationszeit einer Infektion berücksichtigt werden, um die gemeldeten Fallzahlen mit den Infektionen vergleichen zu können.

Gründe für den Meldeverzug können sein, dass sich Infizierte erst nach Einsetzen der Symptome um einen Test bemühen. Daraufhin folgt die Wartezeit auf einen Termin, die Durchführung und Evaluierung der Tests und die Bekanntmachung und Veröffentlichung der Testergebnisse. Bei der Berechnung der Dunkelziffer werden zunächst adjustierte Infektionszahlen  $c'_{t,m}$  berechnet, die die Zahl der Infektionen von Land  $m$ , die zum Zeitpunkt  $t$  hätten gemeldet werden müssen, beschreiben. Da der bereitgestellte *reporting delay* erst nach Berücksichtigung der Inkubationszeit dazuaddiert werden darf, werden zunächst analog wie in (2.8) die bekannten Todeszahlen  $d_{t,m}$  herangezogen, um mithilfe der zuvor eingeführten Verteilung  $X_2$ , die Zeit zwischen dem Auftreten von Symptomen und Todeszeitpunkt von den Todeszahlen abzuziehen.  $incub_{t,m}$  bezeichnet somit die erwartete Anzahl an Infizierten in dem Land  $m$ , die an ihrer Infektion sterben werden und zum Zeitpunkt  $t$  Symptome zeigen. Schließlich muss auf die  $incub_{t,m}$  der *reporting delay* dazuaddiert werden. Die bereitgestellten Daten zeigen, dass  $\sum_{i=1}^{61} P(\text{reporting delay} = i) = 1$ , und somit gilt für alle  $i > 61$  die Wahrscheinlichkeit  $P(\text{reporting delay} = i) = 0$ . Das bedeutet, dass es nie zu einer Verzögerung von über 61 Tagen hinaus kommt. Natürlich werden nicht alle Menschen die infiziert sind auch sterben, deshalb muss abschließend noch durch die *ifr* dividiert werden:

$$\begin{aligned} E(X_2) &= 17.8, \quad Var(X_2) = 64.1601 \\ \Rightarrow X_2 &\sim \text{Gamma}(4.9383, 0.2774) \\ incub_{t,m} &= \sum_{i=1}^{60} d_{t+i,m} \cdot (P(X_2 \leq i + 0.5) - P(X_2 \leq i - 0.5)) \end{aligned}$$

$$c'_{t,m} = \frac{\sum_{i=1}^{61} incub_{t-i,m} \cdot P(\text{reporting delay} = i)}{ifr} \quad (2.10)$$

Letzten Endes darf nicht vergessen werden, dass die Infektionszahlen aus den gemeldeten Todesfällen geschätzt wurden. Es ist davon auszugehen, dass auch die Todesfälle, ähnlich wie die Infektionen, einen Meldeverzug aufweisen. Es liegen für diese Arbeit keine Daten über den Meldeverzug der durch eine SARS-CoV-2-Infektion verursachten Todesfälle vor. Die Ergebnisse der Infektionszahlberechnung unter Berücksichtigung des *reporting delays* aus Kapitel (3.2) werden jedoch zeigen, dass es notwendig ist, einen solchen Meldeverzug der Todeszahlen zu berücksichtigen. Anhand dieser Ergebnisse kann geschätzt werden, wie hoch ein mittlerer Meldeverzug der Todesfälle ungefähr sein sollte, sodass die geschätzten Infektionszahlen und die gemeldeten Fälle zueinander passen. Um diesen grob geschätzten, konstanten und mittleren Meldeverzug werden die Infektionszahlen sozusagen wieder zurück in die Vergangenheit verschoben, da eben davon auszugehen ist, dass die Todesfälle um diesen geschätzten Meldeverzug früher aufgetreten sind, als die Daten zu den gemeldeten Todesfällen hergeben.

## 2.3 Theorie der Zeitreihenanalyse

### 2.3.1 Einführung: Das einfache lineare Modell

Bevor im nächsten Abschnitt die Besonderheiten der Regressionsmodelle von Zeitreihen thematisiert werden, wird zunächst ein sehr kurzer Einblick in die lineare Einfachregression als Grundlage der Zeitreihenregressionsmodelle gegeben. Es seien  $n$  Untersuchungseinheiten gegeben, eine Zielvariable  $y^{(i)}$  der Untersuchungseinheit  $i$  wird als von den Einflussvariablen  $x_1^{(i)}$  und  $x_2^{(i)}$  abhängig angenommen. Es gelten dabei folgende Beziehungen:

$$y^{(i)} = \alpha + \beta_1 \cdot x_1^{(i)} + \beta_2 \cdot x_2^{(i)} + \epsilon^{(i)} \quad (2.11)$$

$\epsilon^{(i)}, \dots, \epsilon^{(n)}$  sind unabhängig und identisch verteilt mit:

$$\begin{aligned} E(\epsilon^{(i)}) &= 0 \\ \text{Var}(\epsilon^{(i)}) &= \sigma^2 \end{aligned}$$

Die Aufgaben der statistischen Inferenz sind nun die Schätzungen der Parameter  $\alpha$ ,  $\beta$  und  $\sigma$ . Von besonderem Interesse ist dabei der Parameter  $\beta$ , welcher den Einfluss der Variablen  $x_i$  auf die interessierende Zielgröße  $y_i$  beschreibt. Zur Schätzung der Parameter  $\alpha$  und  $\beta$  wird der gewöhnliche *KQ-(Kleinste-Quadrate)Schätzer* angewandt. Dieser minimiert die Summe der quadratischen Abweichungen zwischen den beobachteten Werten der Zielgröße und den durch das lineare Modell modellierten Werten der Zielgröße, also die Summe

$$\sum_{i=1}^n (y_i - (\alpha + \beta_1 \cdot x_1^{(i)} + \beta_2 \cdot x_2^{(i)}))^2$$

Schreibe man nun

$$\epsilon = \begin{pmatrix} \epsilon^{(1)} \\ \dots \\ \epsilon^{(n)} \end{pmatrix}, Y = \begin{pmatrix} y^{(1)} \\ \dots \\ y^{(n)} \end{pmatrix}, \beta = \begin{pmatrix} \alpha \\ \beta_1 \\ \dots \\ \beta_p \end{pmatrix}$$

bei  $p$  verschiedenen Einflussvariablen und  $X$  als

$$X = \begin{pmatrix} 1 & x_1^{(1)} & \dots & x_p^{(1)} \\ 1 & x_1^{(2)} & \dots & x_p^{(2)} \\ \vdots & \vdots & & \vdots \\ 1 & x_1^{(n)} & \dots & x_p^{(n)} \end{pmatrix},$$

so lautet für die Gleichung  $Y = X \cdot \beta + \epsilon$  die analytische Lösung des *KQ – Schätzers*  $\hat{\beta} = (X^T X)^{-1} X^T Y$ .

### 2.3.2 Zeitreihenanalyse: Statische vs. Dynamische Modelle

Zeitreihen sind Daten einer Untersuchungseinheit über mehrere Zeitpunkte. Wird mehr als eine Größe beobachtet, spricht man von einer multivariaten Zeitreihe. Damit können kausale Zusammenhänge mehrerer Variablen und derer Veränderungen im Zeitverlauf analysiert werden. Regressionsmodelle zur Analyse von Zeitreihendaten können in zwei Kategorien unterteilt werden, abhängig davon, ob die betrachteten Größen, die zeitlich geordnet werden können, miteinander korrelieren oder nicht. Es ist nämlich möglich, dass eine beobachtete Einflussgröße die Zielgröße nicht nur zum gleichen Zeitpunkt beeinflusst, sondern auch noch in den Folgezeitpunkten. Darauf wird im Abschnitt *Dynamische Modelle* näher eingegangen [Stoetzer, 2020].

#### Statische Modelle

Im einfachsten Fall hängen die Beobachtungen einer Variablen nicht von den Beobachtungen der vorherigen bzw. nachfolgenden Zeitpunkte ab. Es genügt in diesem Fall eine einfache lineare Modellspezifikation, welche nur kontemporäre Wirkungen berücksichtigt. Für den Zeitpunkt  $t$  gelten analog zum einfachen linearen Modell (2.11) folgende Gleichungen:

$$y^{(t)} = \alpha + \beta \cdot x^{(t)} + \epsilon^{(t)}$$

alle  $\epsilon^{(t)}$  sind unabhängig und identisch verteilt mit:

$$\begin{aligned} E(\epsilon^{(t)}) &= 0 \\ \text{Var}(\epsilon^{(t)}) &= \sigma^2 \end{aligned}$$

## Dynamische Modelle

Bei Zeitreihendaten sind die Beobachtungen zeitlich sortiert, und es ist wie eingangs erwähnt möglich, dass diese miteinander korrelieren. Veränderungen einer Einflussvariable  $x$  beeinflussen eine Zielgröße  $y$  möglicherweise nicht nur zum Zeitpunkt der Veränderung  $t$ , sondern womöglich auch darüber hinaus in den folgenden Zeitpunkten  $t+1, t+2$  usw. Wenn solche Veränderungen ihre Wirkung über einen längeren Zeitraum entfalten, spricht man von dynamischen Beziehungen. Der Wert der Variablen  $y$  zum Zeitpunkt  $t$  hängt damit nicht nur vom Wert der Variablen  $x$  zum Zeitpunkt  $t$ , sondern auch zum Zeitpunkt  $t-1, t-2$  usw. ab. Beeinflusst eine Variable  $x$  die Zielgröße zu den Zeitpunkten  $t, t+1$  und  $t+2$ , so ergibt sich folgende lineare Spezifikation:

$$y^{(t)} = \alpha + \beta_1 \cdot x^{(t)} + \beta_2 \cdot x^{(t-1)} + \beta_3 \cdot x^{(t-2)} + \epsilon^{(t)} \quad (2.12)$$

Die in der Gleichung (2.12) vorkommenden Einflussvariablen  $x^{(t-1)}$  und  $x^{(t-2)}$  werden als Lag erster Ordnung und Lag zweiter Ordnung der Variablen  $x$  bezeichnet. Nimmt man solche Lags in die Modellgleichung mit auf, so spricht man von sogenannten *Distributed-Lag-Modellen*, also von Modellen mit verteilten Verzögerungen. Die Koeffizienten  $\alpha, \beta_1, \beta_2$  und  $\beta_3$  können daraufhin ganz gewöhnlich wie in (2.3.1) mit dem *KQ-Schätzer* geschätzt werden. Auch kann die Gleichung (2.12) um weitere Einflussvariablen  $x_1, x_2$  usw. ergänzt werden. Diese können auch verzögerte Effekte ausüben, damit können in der Gleichung (2.12) auch weitere Lags  $x_1^{(t-1)}, x_1^{(t-2)}, x_2^{(t-1)}, x_2^{(t-2)}$  usw. aufgenommen werden. Die Anzahl der Lags einzelner Einflussvariablen kann sich selbstverständlich voneinander unterscheiden. Wie viele Lags jeweils in die Modellgleichung miteinbezogen werden, sollte in erster Linie inhaltlich begründet werden. Der langfristige Einfluss der Variablen  $x$  aus (2.12) setzt sich dann zusammen aus  $\beta_1 + \beta_2 + \beta_3$ , sofern  $x$  keinen weiteren Einfluss über diese drei Zeitpunkte hinweg besitzt. Es besteht theoretisch auch die Möglichkeit, sogenannte Leads in die Analyse miteinzubeziehen. Dies ist dann notwendig, wenn zukünftige Veränderungen einer Einflussvariablen bereits in der Gegenwart einen Einfluss auf die Zielvariable ausüben. In dem Fall würde das Modell aus (2.12) um die Einflussvariablen  $x^{(t+1)}, x^{(t+2)}$  usw. ergänzt. Auf diese Möglichkeit wird in dieser Arbeit aufgrund mangelnder inhaltlicher Notwendigkeit verzichtet. Auch auf die Einbeziehung sogenannter autoregressiver Effekte, bei denen der Wert der Zielvariablen  $y$  zu einem früheren Zeitpunkt  $t-1, t-2$  usw. als Einflussvariable in die Gleichung (2.12) miteinbezogen wird, wird in dieser Arbeit verzichtet, um ein möglichst übersichtliches Modell zu erhalten, das den Einfluss

der Variablen *Testgröße* und *Positivrate* in den Fokus nimmt. Darauf wird später in den Kapiteln (3.4) und (3.6) näher eingegangen.

## 2.4 Evaluationen

### 2.4.1 Teil I: Autokorrelation der Fehler

Es gibt zwei grundsätzliche Annahmen bei der Regressionsanalyse von Zeitreihen. Zum einen muss Stationarität bei den betrachteten Variablen vorliegen, zum anderen darf keine Autokorrelation der Fehler vorliegen. Das Problem der Stationarität wird im nächsten Kapitel (2.4.2) behandelt. Der Umgang mit der Autokorrelation der Residuen hingegen, wird im Folgenden behandelt. Eine Voraussetzung für die *KQ-Methode* bei Zeitreihen besteht darin, dass keine Autokorrelation der Fehler vorliegen darf. Von Autokorrelation spricht man, wenn eine Variable zum Zeitpunkt  $t$  mit ihren eigenen Werten in den davorliegenden Zeitpunkten korreliert. Die Korrelation mit dem unmittelbar vorangegangenen Zeitpunkt bezeichnet man als Autokorrelation erster Ordnung, diejenige mit früheren Zeitpunkten heißen entsprechend Autokorrelationen höherer Ordnung. Die Autokorrelation erster Ordnung einer Variablen  $x$  ist somit folgendermaßen definiert:

$$\text{KorrK}(x^{(t)}, x^{(t-1)}) = \frac{\text{Cov}(x^{(t)}, x^{(t-1)})}{\sqrt{\text{Var}(x^{(t)})} \cdot \sqrt{\text{Var}(x^{(t-1)})}} \quad (2.13)$$

Autokorrelationen höherer Ordnungen, z.Bsp. zweiter oder dritter Ordnung, erhält man, indem man in der Gleichung (2.13)  $x^{(t-1)}$  jeweils durch  $x^{(t-2)}$  bzw. durch  $x^{(t-3)}$  ersetzt. Autokorrelationen können entsprechend gewöhnlicher Korrelationskoeffizienten nach Pearson, Werte zwischen minus eins und plus eins annehmen. Hierbei repräsentiert ein Wert von plus eins eine perfekte positive lineare Korrelation beider Variablen und ein Wert von minus eins eine perfekte negative Korrelation. Werte nahe null sprechen gegen eine lineare Korrelation beider betrachteter Variablen. Bei der Analyse von Zeitreihen stellen autokorrelierte Einflussvariablen kein Problem bei der Regressions-schätzung dar. Probleme tauchen nur dann auf, wenn die Fehlerterme autokorreliert sind. Dies ist meistens dann der Fall, wenn die Vergangenheit einen Einfluss auf die Gegenwart und Zukunft hat und diese Einflüsse nicht in die Regressionsgleichung mit aufgenommen wurden. Autokorrelierte Fehler zeichnen sich dadurch aus, dass bei einem großen Fehler zum Zeitpunkt  $t$ , Fehler naheliegender vorangegangener oder nachfolgender Zeitpunkte ebenfalls dazu tendieren, groß auszufallen (natürlich genau dann, wenn die Autokorrelation positiv ausfällt). Eine Autokorrelation der Fehler  $\epsilon$  kann durch die Residuen  $\hat{\epsilon}$ , also der Differenz aus den Werten der Zielvariablen und ihrer geschätzten

Werte, geschätzt werden. Es gilt:

$$\hat{\epsilon}^{(t)} = y^{(t)} - \hat{y}^{(t)}$$

wobei

$$\hat{y}^{(t)} = \hat{\alpha} + \hat{\beta} \cdot x^{(t)}$$

Dabei stellen  $\hat{\alpha}$  und  $\hat{\beta}$  die durch die *KQ-Methode* geschätzten Koeffizienten der Regressionsgleichung dar. Unter der Annahme einer korrekten Spezifikation der Regressionsgleichung, führt eine Autokorrelation der Fehler bei einer *KQ-Schätzung* zwar zu unverzerrten Schätzungen des Parameter  $\beta$ , allerdings sind die Standardabweichungen dieser Koeffizienten verzerrt. Dies führt zu unzuverlässigen t-Tests und F-Tests, welche auf diese Standardabweichungen basieren.

Zur visuellen Inspektion von Autokorrelation dienen Autokorrelationsfunktionen (Autocorrelation Function ACF). Diese tragen auf ihrer x-Achse die Lags der Residuen in aufsteigender Reihenfolge und auf der y-Achse die berechnete Autokorrelation ab.

Zusätzlich können die geschätzten Modelle auch mithilfe statistischer Tests auf Autokorrelation überprüft werden. Eine Möglichkeit dafür bietet der *Breusch-Godfrey-Test*. Dieser Test berücksichtigt zwei grundlegende Aspekte. Zum einen lässt er eine Abhängigkeit der Residuen  $\hat{\epsilon}^{(t)}$  von der Zeit  $t$  zu, und zum anderen eine Abhängigkeit der Residuen zum Zeitpunkt  $t$  mit ihrem Wert zum vorangegangenen Zeitpunkt  $t - 1$ . Es ergibt sich die Gleichung:

$$\hat{\epsilon}^{(t)} = \alpha + \beta_1 \cdot t + \beta_2 \cdot \hat{\epsilon}^{(t-1)} + u_t \tag{2.14}$$

$u_t$  ist dabei ein neuer Fehlerterm, welcher einer unabhängigen und identischen Verteilung folgt. Die Nullhypothese lautet nun, dass  $\beta_2 = 0$ . Ist dies nicht der Fall, liegt Autokorrelation erster Ordnung vor. Die Gleichung (2.14) kann auch erweitert werden um somit auch auf Autokorrelationen höherer Ordnung testen zu können. In unserem Fall überprüfen wir die Autokorrelation bis zur zehnten Ordnung. Das heißt (2.14) wird erweitert zu:

$$\hat{\epsilon}^{(t)} = \alpha + \beta_1 \cdot t + \beta_2 \cdot \hat{\epsilon}^{(t-1)} + \beta_3 \cdot \hat{\epsilon}^{(t-2)} + \beta_4 \cdot \hat{\epsilon}^{(t-3)} + \dots + \beta_{11} \cdot \hat{\epsilon}^{(t-10)} + u_t$$

Der Test analysiert nun, ob irgendeiner der Koeffizienten von  $\beta_2$  bis  $\beta_{11}$  der Regression sich von 0 unterscheidet. Er verwendet dabei eine *Lagrange-Multiplikator-Statistik (LM-Statistik)* als Testgröße und wird deshalb auch als *Lagrange-Multiplier-Test* bezeichnet. Eine Möglichkeit mit Autokorrelation der Residuen umzugehen ist, diese durch eine Neuspezifikation des Regressionsmodells zu eliminieren. Dies erfolgt typischerweise durch die Einbindung von Lags der Einflussvariablen. Diese Möglichkeit wird in Kapitel (3.6) umgesetzt.

Eine zweite Möglichkeit des Umgangs mit dem Vorliegen von Autokorrelation der Residuen stellt die Neuschätzung der Standardabweichungen der geschätzten Koeffizienten dar. Da, wie eingangs erwähnt, nur diese bei vorhandener Autokorrelation der Residuen verzerrt sind, nicht jedoch die Parameterschätzungen an sich, reicht es, nur diese zu korrigieren. Dafür eignet sich die Schätzung von *Newey-West-Standardfehlern*, welcher die Verzerrungen aufgrund von Autokorrelation eliminiert und zudem auch robust gegen Heteroskedastie, also eine nicht-konstante Varianz der Fehler, ist. Normalerweise würden die Varianzen des gewöhnlichen *KQ-Schätzers* mit folgender Formel geschätzt werden :

$$\text{Var}(\hat{\beta}) = \sigma^2 \cdot (X^T X)^{-1} \quad (2.15)$$

Ein gegen Autokorrelation robuster Schätzer hingegen bezieht die wachsende Korrelation der Fehler bei Verkleinern des Zeitabstands zwischen den Fehlern mit ein. Man erhält:

$$\text{Var}^{\text{unverzerrt}}(\hat{\beta}) = (X^T X)^{-1} \cdot X^T S X \cdot (X^T X)^{-1} \quad (2.16)$$

Wobei für  $X^T S X$  folgendes gilt:

$$X^T S X = \frac{1}{n} \sum_{t=1}^n \epsilon_t^2 \cdot (X^T X) + \frac{1}{n} \sum_{l=1}^L \sum_{t=l+1}^n w_l \cdot \epsilon_t \cdot \epsilon_{t-l} \cdot (X_t^T X_{t-l} + X_{t-l}^T X_t) \quad (2.17)$$

$$w_l = 1 - \frac{l}{L+1}$$

$w_l$  können als Gewichte angesehen werden, die zeitlich näher beieinanderliegenden Datenpaaren höher wertet.  $n$  ist die Anzahl aller vorhandenen Zeitpunkte aus dem Datensatz und  $L$  ist die Zahl der Lags, bis zu der Autokorrelation der Residuen vermutet wird. Als Faustregel wird dabei gemeinhin angenommen, dass  $L = 0.75 \cdot n^{\frac{1}{3}}$  [Stock und Watson, 2015].

## 2.4.2 Teil II: Stationarität der Variablen

Neben dem Ausschließen von Autokorrelation der Fehler muss noch die Stationarität aller gegebenen Einfluss- und Zielvariablen gegeben sein. Stationarität bedeutet einfach ausgedrückt, dass die Struktur einer Zeitreihe sich in der Zukunft nicht von der Vergangenheit unterscheiden darf. Sie liegt dann vor, wenn die Varianz und der Erwartungswert aller Variablen im betrachteten Zeitraum konstant bleiben. Somit ist diese nicht gegeben, wenn bspw. Trends, saisonale Effekte oder Strukturbrüche vorliegen. Bei Nicht-Stationarität besteht insbesondere die Gefahr, Scheinkorrelationen zu modellieren ([S.136ff.][Stoetzer, 2020]).

Kapitel (3.5.2) wird zeigen, dass alle unserer betrachteten Variablen nicht-stationär sind. Dies führt zu Problemen bei der Zeitreihenanalyse. Eine Möglichkeit, dieses Problem zu lösen, ist die Bildung von Differenzen unserer Variablen [S.149][Stoetzer, 2020]. Erste Differenzen werden gebildet, indem die Differenz einer Variablen zwischen zwei aufeinanderfolgenden Zeitpunkten berechnet wird. Es wird also vom Wert einer Variablen zum Zeitpunkt  $t$  der Wert dieser Variable zum vorangegangenen Zeitpunkt  $t - 1$  abgezogen. Erste Differenzen sind kurzfristige Veränderungen, langfristige Zusammenhänge werden dabei eliminiert. Diese Analyse von Veränderungen ist bei Zeitreihen häufig ein wichtiger Aspekt[Stoetzer, 2020].

Unterscheiden sich Erwartungswerte und Varianzen der Einfluss- und Zielvariablen in verschiedenen Zeiträumen grundlegend, spricht man von sogenannten Strukturbrüchen. Eine Möglichkeit mit so einer Datenlage umzugehen besteht darin, sogenannte *Dummy-Variablen* zur Unterscheidung verschiedener zeitlicher Phasen einzuführen. Die dazugehörige Modellgleichung integriert dafür *Dummy-Variablen*, die nur die Werte 0 oder 1 annehmen können. Mit deren Hilfe können mehrere zeitliche Abschnitte vorliegender Daten getrennt betrachtet werden.

## 2.5 Aufstellung der Modellgleichungen

Auf Basis der theoretischen Annahmen der letzten Kapitel werden nun Modellgleichungen aufgestellt, um die Zusammenhänge zwischen Dunkelziffer, Testgröße und Testpositivrate zu verstehen. Um die Intentionen hinter den in Abschnitt (2.5.2) aufgestellten Modellgleichungen zu verstehen, werden zunächst im nächsten Abschnitt grundlegende Überlegungen zu unseren Variablen aufgezeigt. Dafür wird sofern es nötig ist, auf Ergebnisse vorgegriffen, die erst im nächsten großen Kapitel (3) genau analysiert werden.

### 2.5.1 Problematiken des Zusammenhangs zwischen Infektionen, Dunkelziffer, Positivrate und Testgröße

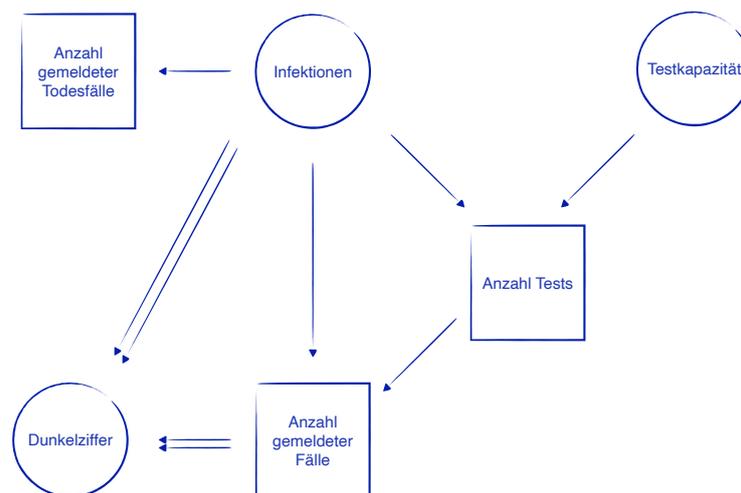


Abbildung 2.1: Flussdiagramm zu den Zusammenhängen zwischen Infektionen, Dunkelziffer, Positivrate und Testgröße

Abbildung (2.1) zeigt die Hauptproblematik auf, mit der man in dieser Arbeit konfrontiert ist. Es wird sich in Kapitel (3) zeigen, dass es bei der Analyse des Zusammenhangs zwischen der Dunkelziffer auf der einen Seite, und der Testgröße und der Testpositivrate auf der anderen Seite, einige schwerwiegende Hindernisse gibt, die sich mit dem Flussdiagramm aus Abbildung (2.1) gut erklären lassen. Diese Abbildung zeigt, wie sich die Zahl der Todesfälle, der Infektionen, der Testgröße, und der Dunkelziffer, die Testkapazität und die Testpositivrate gegenseitig beeinflussen. Natürlich beeinflusst die Zahl der

Infektionen die Zahl der Todesfälle, deswegen werden in dieser Arbeit auch die Zahl der Todesfälle genutzt, um die Zahl der Infektionen zu schätzen. Die Testkapazität, also das Angebot an durchführbaren Tests, beeinflusst die Zahl der tatsächlich durchgeführten Tests. Diese Kapazität ist allerdings nicht berücksichtigt in dieser Arbeit. Die Zahl der durchgeführten Tests und die Zahl der Infektionen beeinflussen letztendlich die Zahl der gemeldeten Fälle. Aus den gemeldeten Fällen und der Testgröße lässt sich deterministisch die Testpositivrate berechnen und aus der Zahl der gemeldeten Fälle und der Infektionen lässt sich schließlich auch deterministisch die Dunkelziffer berechnen. Wären dies die einzigen bedeutungsvollen Zusammenhänge zwischen diesen betrachteten Größen, könnte man einfacher den Einfluss der Testgröße und der Testpositivrate auf die Dunkelziffer schätzen. Genauere Analysen der vorliegenden Zahlen in Kapitel (3) werden jedoch zeigen, dass dies nicht der Fall ist. Es wird sich in dieser Arbeit zeigen, dass zusätzlich die Zahl der Infektionen die Zahl der durchgeführten Tests und somit auch die Testpositivrate sehr stark beeinflusst. Im Verlauf der Pandemie steigt die Zahl der durchgeführten Tests in Deutschland und auch in den anderen vier betrachteten Ländern sehr deutlich. Gründe dafür können sein, dass politische Anstrengungen zur Eindämmung der Pandemie die Testkapazität im Verlauf der Monate stetig erhöht, und dass die Kopplung von Freiheiten an negativen Tests die Zahl der durchgeführten Tests zusätzlich erhöht. Gleichzeitig kann sich das Coronavirus seit Beginn des Jahres 2020 immer weiter ausbreiten, die Infektionszahlen steigen auch deutlich im Verlauf der Monate, wenn auch eher in immer wiederkehrenden Wellen mit zwischenzeitlichen Rückgängen der Infektionszahlen. Auf alle diese Entwicklungen wird allerdings erst im nächsten Kapitel (3) genauer eingegangen. Die stetige Ausbreitung von SARS-CoV-2 ruft also als Reaktion eine Erhöhung der durchgeführten Tests hervor, somit beeinflusst die Zahl der Infektionen die Testgröße. Möchte man nun jedoch den Einfluss der Testgröße auf die Dunkelziffer bestimmen, ergibt sich daraus ein Problem. Ein weiteres Problem stellt die Tatsache dar, dass die Zahl der Infektionen auch die Dunkelziffer stark beeinflusst, denn eine starke Zunahme der Infektionen führt auch automatisch zu einer starken Zunahme der Dunkelziffer. Auch darauf wird im nächsten Kapitel genauer eingegangen. Dieser direkte Einfluss der Infektionen auf Dunkelziffer, die Testgröße und die Positivrate stellt das Kernproblem dieser gesamten Arbeit dar. Sie ist auch die Basis der unterschiedlichen Ideen zur Modellierung des Zusammenhanges zwischen Dunkelziffer, Testgröße und Positivrate, was in einer Aufstellung mehrerer verschiedener Modelle resultiert. Diese werden im nächsten Abschnitt beleuchtet.

## 2.5.2 Modellgleichungen

### Statische Modelle

Zur Modellierung des Zusammenhangs zwischen Testgröße, Positivrate und Dunkelziffer in Deutschland, werden zunächst insgesamt fünf verschiedene statische Modelle aufgestellt. Das erste Modell

$$\text{Dunkelziffer} = \alpha + \beta_1 \cdot \text{Tests} + \beta_2 \cdot \text{Positivrate} + \epsilon$$

zeigt eine erste naive Herangehensweise an die Problemstellung. Es modelliert den Effekt der Anzahl der Tests pro 100000 Einwohner und der Positivrate auf die absolute Dunkelziffer. Die Ergebnisse in Kapitel (3.4.1) werden zeigen, dass ein solches Modell für unsere Datenlage nicht geeignet ist. Sie führen zu Überlegungen über eine getrennte Betrachtung des Einflusses der Testgröße und des Einflusses der Positivrate. Zwei verschiedene Modelle zur Betrachtung des Einflusses der Testgröße werden aufgestellt,

$$\text{Anteil der Dunkelziffer} = \alpha + \beta_1 \cdot \text{Tests} + \epsilon$$

$$\text{und } \text{Anteil der Dunkelziffer} = \alpha + \beta_1 \cdot \log(\text{Anteil der Tests}) + \epsilon.$$

Beide Modelle haben als Zielgröße den *Anteil der Dunkelziffer*, welche den Anteil der Dunkelziffer an der Gesamtzahl der Infektionen beschreibt. Das erste Modell hat immer noch die Anzahl der Tests pro 100000 Einwohner als Einflussgröße, während das zweite auch hier den Anteil an der Gesamtzahl der Infektionen betrachtet. Zwei verschiedene Modelle werden zur Modellierung des Einflusses der Positivrate aufgestellt. Das erste Modell

$$\text{Dunkelziffer} = \alpha + \beta_1 \cdot \text{Positivrate} + \epsilon$$

modelliert dabei den Zusammenhang zwischen der Positivrate und der Dunkelziffer.

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \epsilon$$

beschreibt abschließend den direkten Zusammenhang zwischen der Positivrate und den Infektionszahlen.

### Modell unter Differenzbildung

Die letzten drei statischen Modellgleichungen, die in dieser Arbeit betrachtet werden, entstehen aus der Intention heraus, eine Lösung im Umgang mit der Nicht-Stationarität unserer Variablen zu finden. Man erhält unter der Berücksichtigung der Möglichkeit von Differenzbildungen folgende Modellgleichungen:

$$\begin{aligned} \text{Diff}(\text{Anteil der Dunkelziffer}) &= \alpha + \beta_1 \cdot \text{Diff}(\log(\text{Anteil der Tests})) + \epsilon \\ \text{und } \text{Diff}(\text{Infektionen}) &= \alpha + \beta_1 \cdot \text{Diff}(\text{Positivrate}) + \epsilon \end{aligned}$$

### Modell unter Berücksichtigung von Strukturbrüchen

Ebenso wie die beiden *Modelle unter Differenzbildung* versucht auch die letzte Modellgleichung die Problematik der Nicht-Stationarität einzufangen. Dafür betrachtet sie separat unterschiedliche Phasen der Pandemie. Für Deutschland ergibt sich dabei die getrennte Betrachtung dreier zeitlicher Abschnitte, wie Kapitel (3.5.2) zeigen wird. Zwei davon beschreiben Infektionswellen, die dritte eine ruhigere Phase zwischen diesen beiden Infektionswellen. Es ergibt sich die Modellgleichung:

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \beta_2 \cdot \text{erste Welle} + \beta_3 \cdot \text{zweite Welle} + \epsilon$$

### Dynamische Modelle

Zur Modellierung des Zusammenhangs zwischen Testgröße, Positivrate und Dunkelziffer werden in Kapitel (3.6) zwei dynamische Modelle aufgestellt und berechnet. Dafür werden die beiden statischen Modelle, welche sich in Kapitel (3.4) als am geeignetsten herausgestellt haben, durch die Integration von Lags in dynamische Modelle überführt. Aus dem statischen Modell:

$$\text{Anteil der Dunkelziffer} = \alpha + \beta_1 \cdot \log(\text{Anteil der Tests}) + \epsilon,$$

wird das dynamische Modell:

$$\begin{aligned} \text{Anteil der Dunkelziffer}^{(t)} &= \\ &\alpha + \beta_1 \cdot \log(\text{Anteil der Tests})^{(t)} + \beta_2 \cdot \log(\text{Anteil der Tests})^{(t-1)} + \\ &\beta_3 \cdot \log(\text{Anteil der Tests})^{(t-2)} + \beta_4 \cdot \log(\text{Anteil der Tests})^{(t-3)} + \\ &\beta_5 \cdot \log(\text{Anteil der Tests})^{(t-4)} + \beta_6 \cdot \log(\text{Anteil der Tests})^{(t-5)} + \epsilon^{(t)}. \end{aligned}$$

Und aus dem statischen Modell:

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \epsilon,$$

wird das dynamische Modell:

$$\begin{aligned} \text{Infektionen}^{(t)} = & \\ & \alpha + \beta_1 \cdot \text{Positivrate}^{(t)} + \beta_2 \cdot \text{Positivrate}^{(t-1)} + \beta_3 \cdot \text{Positivrate}^{(t-2)} + \\ & \beta_4 \cdot \text{Positivrate}^{(t-3)} + \epsilon^{(t-2)}. \end{aligned}$$

Im übernächsten Kapitel werden all diese vorgestellten Modelle berechnet, diskutiert und evaluiert. Im nächsten Abschnitt (2.6) folgt jedoch ein kurzer theoretischer Exkurs zu der Rolle von Sensitivität und Spezifität bei der Durchführung von medizinischen Tests. Ergebnisse dazu mit unseren Daten werden erst im letzten Teil dieser Arbeit, dem Ausblick (4), gegeben.

## 2.6 Exkurs: Spezifität und Sensitivität

### 2.6.1 Einführung und Definitionen

Falsch positive und falsch negative Testergebnisse können die Zahl gemeldeter Fälle stark beeinflussen und verzerren. Was das genau für die Zahl nicht entdeckter Fälle bedeuten kann, und wie genau die Zahl der gemeldeten Fälle manipuliert werden kann, wird im letzten Kapitel (4.2) aufgezeigt. Hier werden zunächst nur die theoretischen Hintergründe zu den Faktoren *Spezifität* und *Sensitivität* thematisiert. Methoden hierzu wurden aus Günther et al. [2021] entnommen. Sie ermöglichen die Adjustierung gemeldeter Fallzahlen und die Schätzung einer von Missklassifikationseffekten bereinigten epidemischen Kurve.

Die Spezifität eines Tests ist definiert als die Wahrscheinlichkeit, bei einer gesunden Person ein negatives Testergebnis zu erhalten und entspricht somit dem Quotienten aus richtig negativen Testergebnissen und der Anzahl aller getesteter gesunder Personen. Die Sensitivität eines Tests ist hingegen definiert als die Wahrscheinlichkeit, bei einer erkrankten Person ein positives Testergebnis zu erhalten und entspricht somit dem Quotienten aus richtig positiven Testergebnissen und der Anzahl aller getesteter erkrankter Personen [S.46][Schwartz et al., 1984].

Die Missklassifikation bei der Diagnose von Infektionen mit dem SARS-CoV-2 manifestiert sich also in zweierlei Hinsicht. Zum einen kann einer tatsächlich infizierten Personen fälschlicherweise ein negativer Befund attestiert werden (falsch-negatives Ergebnis), zum anderen kann eine nicht infizierte Person ein positives Testergebnis erhalten (falsch-positives Ergebnis). Beides führt zu Verzerrungen der epidemischen Kurve, wobei falsch-negative Ergebnisse die Zahl der Infizierten nach unten und falsch-positive Ergebnisse nach oben verzerren [Günther et al., 2021].

### 2.6.2 Methode zur Adjustierung der Fallzahlen

Die Art der Adjustierung der Fallzahlen wird im Folgenden beschrieben. Sei hierbei  $NT_t$  die Anzahl getesteter Personen, deren Testergebnis zum Zeitpunkt  $t$  gemeldet wird.  $T_t^+$  sei die Anzahl der Personen mit einem positiven Testergebnis, welches zum Zeitpunkt  $t$  gemeldet wird.  $D_t^+$  sei die Anzahl der tatsächlich Infizierten und Getesteten, deren

positives Ergebnis hätte zum Zeitpunkt  $t$  gemeldet werden müssen. Es sei zu beachten, dass ebendieser Parameter  $D_t^+$  in der Praxis nicht bekannt ist.  $D_t^+$  kann jedoch approximativ berechnet werden unter Zuhilfenahme der bekannten Parameter  $NT_t$  und  $T_t^+$  und unter verschiedenen Annahmen über die Spezifität und Sensitivität der durchgeführten Tests. Unter den Annahmen aus (2.6.1) ergibt sich für den Erwartungswert der positiven Testergebnisse:

$$E(T_t^+ | D_t^+, NT_t) = D_t^+ \cdot \text{Sensitivität} + (NT_t - D_t^+) \cdot (1 - \text{Spezifität}) \quad (2.18)$$

Diese Gleichung beschreibt sehr deutlich die Wirkungsweise von Werten für eine Spezifität und Sensitivität ungleich eins. Der erste Summand bildet dabei die Anzahl richtig-positiver Testergebnisse. Der Einfluss von Werten für eine Sensitivität ungleich eins hängt also nur von der Anzahl tatsächlich Infizierter ab. Eine hohe Inzidenz erhöht also den Effekt der Sensitivität und damit die Wahrscheinlichkeit, die tatsächlichen Fallzahlen zu unterschätzen. Der Einfluss von Werten für eine Spezifität ungleich eins hingegen wirkt anders. Der zweite Summand aus (2.18) bildet die Anzahl falsch-positiver Testergebnisse. Der Effekt von Werten für eine Spezifität ungleich eins auf die gemeldeten Fallzahlen hängt sowohl von der Anzahl tatsächlich Infizierter als auch von der Anzahl der durchgeführten Tests ab. Eine hohe Differenz zwischen durchgeführten Tests und der Anzahl Infizierter unter diesen Getesteten, verstärkt den Einfluss der Spezifität. Somit steigt die Wahrscheinlichkeit, die Fallzahlen zu überschätzen, bei hoher Testaktivität und gleichzeitig niedriger Inzidenz. Sind nun die Testgröße und die Zahl gemeldeter Fälle gegeben, kann die Zahl tatsächlich Infizierter aus (2.18) geschätzt werden und man erhält mittels einfacher Umformungen die dafür benötigte Formel (2.19).

$$\begin{aligned} T_t^+ &= \hat{D}_t^+ \cdot \text{Sens} + NT_t^+ - \hat{D}_t^+ - NT_t \cdot \text{Spez} + \hat{D}_t^+ \cdot \text{Spez} \\ \Leftrightarrow -\hat{D}_t^+ \cdot \text{Sens} + \hat{D}_t^+ - \hat{D}_t^+ \cdot \text{Spez} &= -T_t^+ + NT_t - NT_t \cdot \text{Spez} \\ \Leftrightarrow \hat{D}_t^+ \cdot (-\text{Sens} + 1 - \text{Spez}) &= -T_t^+ + NT_t - NT_t \cdot \text{Spez} \\ \Leftrightarrow \hat{D}_t^+ &= \frac{NT_t \cdot (1 - \text{Spez}) - T_t^+}{1 - \text{Spez} - \text{Sens}} \\ \Leftrightarrow \hat{D}_t^+ &= \frac{T_t^+ - NT_t \cdot (1 - \text{Spez})}{\text{Sens} + \text{Spez} - 1} \end{aligned} \quad (2.19)$$

Die Größen  $T_t^+$  und  $NT_t^+$  sind in dem in dieser Arbeit benutzten Datensatz gegeben. Werte für die Sensitivität und die Spezifität für SARS-CoV-2-PCR-Tests werden

als sehr hoch angenommen. Untersuchungen deutscher Labore ergaben eine durchschnittliche Spezifität zwischen 97.8% und 98.6% und eine durchschnittliche Sensitivität zwischen 98.9% und 99.7% [Günther et al., 2021]. Diese Werte sind jedoch überoptimistisch, da sie nur die analytischen Werte für Sensitivität und Spezifität unter Laborbedingungen widerspiegeln. In der Realität führen viele Faktoren jedoch zu weit geringeren Annahmen über die Spezifität und Sensitivität von durchgeführten SARS-CoV-2-PCR-Tests [Watson et al., 2020].

Bspw. können Verunreinigungen der Proben etwa beim Transport oder ein falsch gewählter Zeitpunkt der Testdurchführung die Sensitivität erheblich verringern und zu falsch-negativen Ergebnissen führen. Aufgrund der noch zu geringen Virenlast bei frisch infizierten Personen liegt die Sensitivität direkt nach einer Infektion nahe bei null. Erst einige Tage später, nachdem zumal auch Symptome einsetzen, erhöht sich die Sensitivität auf ca. 80-90% [Günther et al., 2021]. Analog zu [Günther et al., 2021], werden auch in dieser Arbeit Werte für die Sensitivität von 70% und 90% angenommen. In Anbetracht der Spezifität können gegenseitige Verunreinigungen der Proben zu falsch-positiven Ergebnissen führen und die Spezifität verringern. Auch hier werden analog zu [Günther et al., 2021] Werte für die Spezifität von 99.5%, 99.7% und 99.9% angenommen.

# 3 Ergebnisse

## 3.1 Deskriptive Analyse von Testgröße, gemeldeten Fällen und Positivrate

Abbildung (3.1) zeigt den Verlauf der Positivrate zwischen März 2020 und Februar 2021 in Deutschland. Dabei nimmt sie Werte von um die ein Prozent bis ca. fünfzehn Prozent an. Sie startet bei ungefähr einem Prozent, steigt dann steil an bis sie in der Woche vom fünften April einen Hochpunkt von ca. neun Prozent erreicht. Danach fällt sie wieder steil ab und verweilt zwischen den Monaten Mai und Anfang Oktober bei Werten um die ein Prozent. Ein zweites Mal steigt sie in den Monaten zwischen Oktober 2020 und Anfang Januar 2021 steil an und erreicht einen neuen Höchstwert von fast fünfzehn Prozent. Dabei stagniert sie etwas zwischen den Monaten November und Dezember. Anschließend sinkt sie wieder bis zur Woche vom 14. Februar 2021 auf ein Niveau von ca. fünf Prozent, welches sie in den letzten beiden Wochen relativ konstant hält.

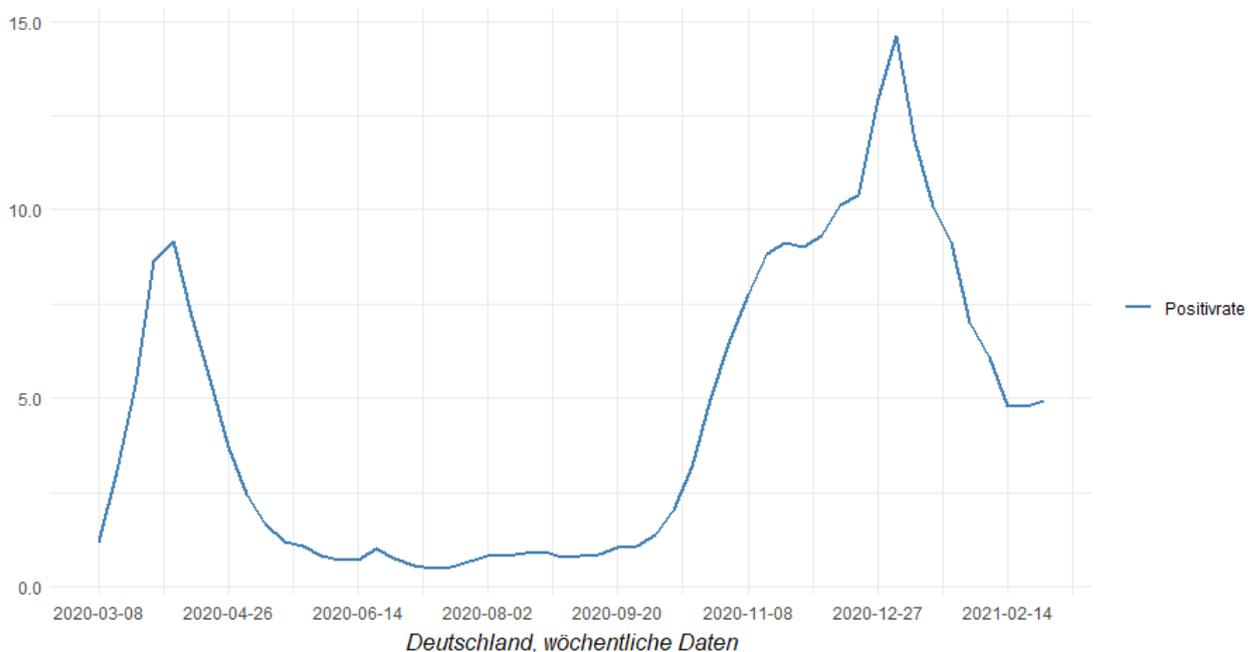


Abbildung 3.1: Positivrate in Deutschland

Abbildung (3.2) zeigt den Verlauf der Anzahl durchgeführter PCR-Tests in Deutschland in dem genau gleichen Zeitraum. Die Testgröße nimmt dabei Werte zwischen

60000 wöchentlich durchgeführter Tests Anfang März 2020 und bis zu 1700000 Tests im Dezember 2020 an. Ein Trend zur Zunahme durchgeführter Tests ist deutlich erkennbar, zumindest bis zu der Woche vom 20. Dezember 2020. Gleich zu Beginn steigt die Testgröße stark an, stagniert dann ungefähr drei Wochen auf einem Niveau von rund 400000 wöchentlich durchgeführten Tests. Zwischen den Sommermonaten Juni und Ende August steigt sie ein zweites Mal kräftig an auf ein Niveau von 1200000, stagniert dann wieder ein bis zwei Monate, bevor sie wieder steigt und ihren Höchstwert von ca. 1700000 Tests erreicht. Nach einem kurzen Einbruch verweilt die Testgröße schließlich auf einem relativ konstanten hohen Niveau von rund 1100000 wöchentlich durchgeführten Tests.

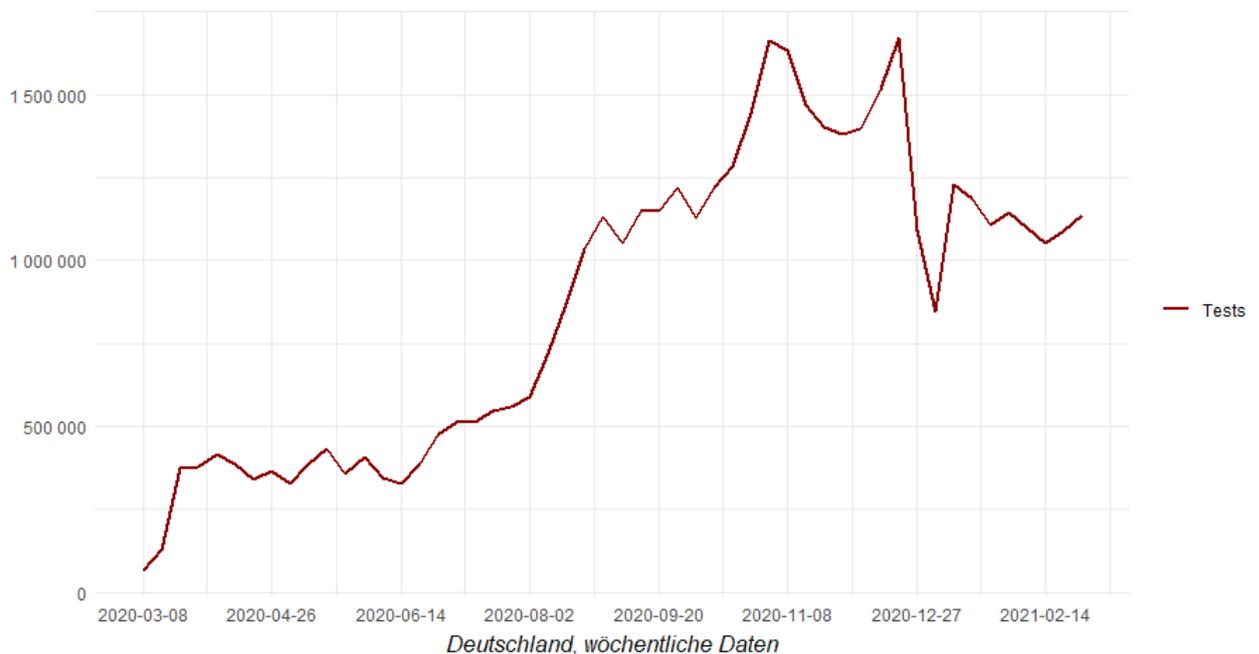


Abbildung 3.2: Anzahl Tests in Deutschland

Vergleicht man die Zahl durchgeführter Tests mit der Anzahl gemeldeter Fallzahlen (Abbildung (3.3)) fällt auf, dass die Zahl durchgeführter Tests die Zahl gemeldeter Fälle um Weiten übertrifft. Dies ergibt Sinn, nachdem die Positivrate wie bereits beschrieben nur niedrige Werte von einem bis maximal fünfzehn Prozent annimmt. Die sichtbaren beiden Wellen erhöhter gemeldeter Fallzahlen zwischen März und April 2020 und zwischen Oktober 2020 und Januar 2021 fallen in dem gleichen Zeitraum wie die beiden Wellen erhöhter Positivraten der Tests. Die deutlich höhere Zahl gemeldeter Fälle in der zweiten Welle verglichen zur ersten Welle trotz ungefähr gleicher Positivraten

lässt sich natürlich durch die erheblich gestiegene Testgröße während der zweiten Welle erklären.

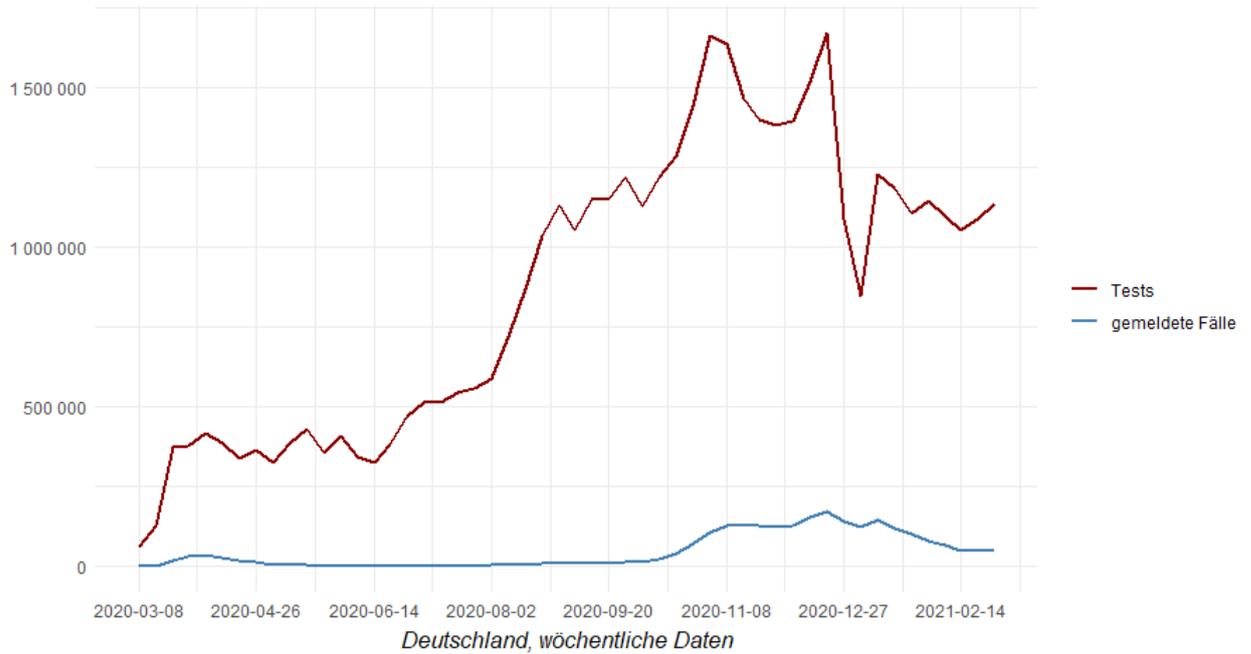


Abbildung 3.3: Tests und gemeldete Fälle in Deutschland

## 3.2 Berechnung der Infektionszahlen

Land	<i>ifr</i>
Deutschland	0,01233
Frankreich	0,01256
Italien	0,01245
Spanien	0,01078
Dänemark	0,01021

Tabelle 3.1: Ergebnisse der *ifr*-Berechnung

Abbildung (3.1) zeigt die Ergebnisse der Berechnung der *infection fatality rates* für die Länder Deutschland, Frankreich, Spanien, Italien und Dänemark. Es zeigt sich, dass die *ifr* mit Werten von 0,01021 bis 0,01256 alle nah beieinander liegen. Das bedeutet, dass geschätzt in diesen Ländern zwischen 1,021% und 1,256% der Infizierten an Corona sterben.

Abbildung (3.4) zeigt die täglich gemeldeten Todesfälle in Deutschland im Zeitraum von Mitte Dezember 2020 bis Ende April 2021. Erste Todesfällen treten im März 2021 auf. Dabei spiegelt die blaue Kurve die tatsächlich gemeldeten Todesfälle wider, während die schwarze Kurve zur besseren Übersicht die über einen Zeitraum von sieben

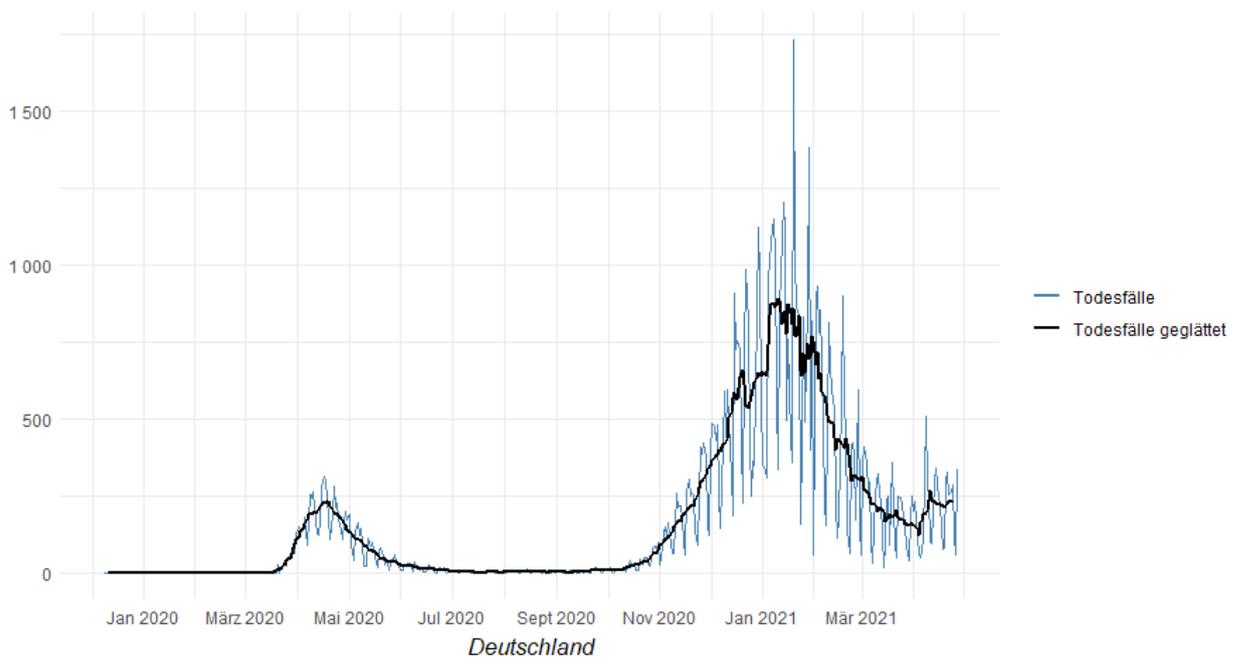


Abbildung 3.4: Todesfälle in Deutschland

Tagen geglätteten Todesfälle zeigt. Diese werden berechnet, indem für jeden Tag der Mittelwert aller Todesfälle der vorangegangenen dreieinhalb Tage, der nachfolgenden dreieinhalb Tage und des betrachteten Tages selber ermittelt wird. Diese Glättung dient lediglich der besseren Visualisierung der Todesfälle, bei der Infektionszahlberechnung werden wie bereits beschrieben die Todesfälle nicht geglättet. Es zeigt sich eine Häufung der durch eine SARS-CoV-2-Infektion verursachten Todesfälle zu zwei Zeitperioden. Die erste Welle an Todesfällen ereignet sich zwischen März 2020 und Juni 2020. Dabei erreicht diese Welle ihren Höhepunkt Mitte April 2020 mit geschätzten täglichen Todesfällen von bis zu 220. Die zweite Welle startet Mitte Oktober 2020, wächst rapide an bis zu einem Höchstwert von fast 900 geschätzten täglichen Todesfällen Mitte Januar 2021. Danach sinken die Zahlen wieder bis zum April 2021. Nach diesem Zeitpunkt kündigt sich auch schon die dritte Welle an neuen Todesfällen an. Die Todesfälle im Sommer 2021, zwischen den Monaten Juni und Oktober, belaufen sich auf Werte von unter zwanzig und befinden sich zum Großteil im einstelligen Bereich.

Abbildung (3.5) verdeutlicht visuell die Berechnung der Infektionszahlen aus den gemeldeten Todeszahlen. Die blaue Kurve stellt wie in Abbildung (3.4) die geglätteten Todeszahlen dar, allerdings wurden diese mit dem Faktor  $\frac{1}{ifr}$  multipliziert. Dies dient ausschließlich visuellen Zwecken, würde man nur die tatsächlichen Todesfälle in dieser

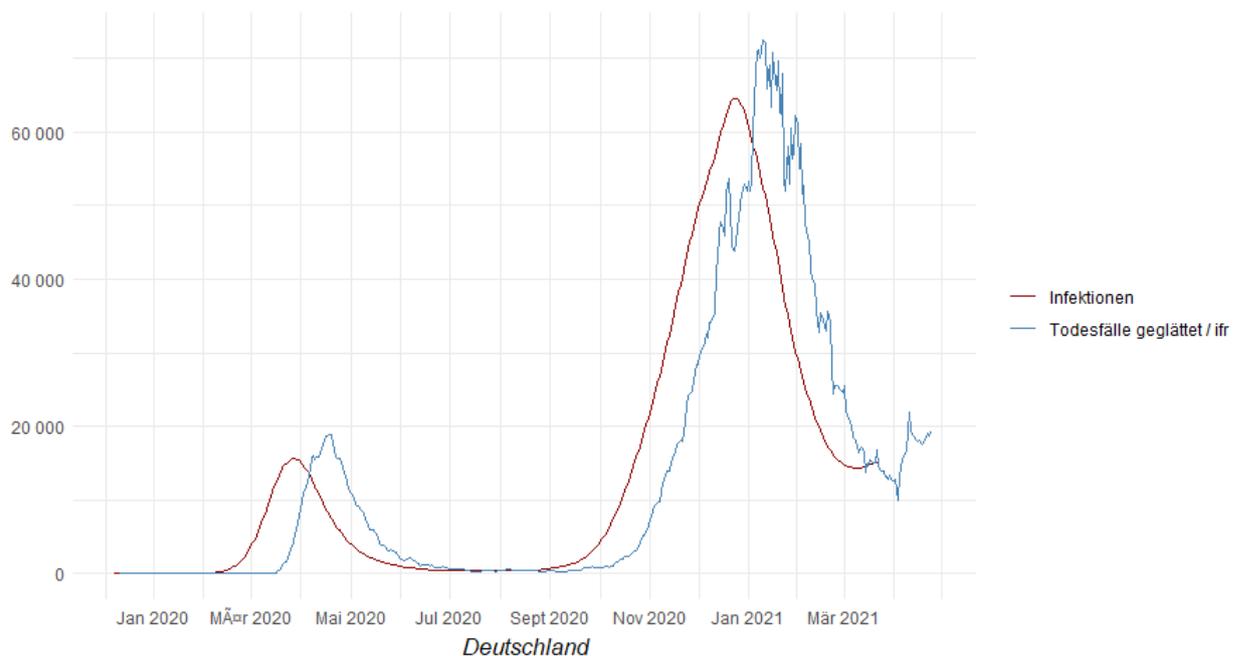


Abbildung 3.5: Modellierte Infektionen in Deutschland

Grafik abbilden, würde man diese aufgrund der Dimension der y-Skala nicht sehen. Es zeigt sich, dass die modellierten Infektionen mit dem SARS-CoV-2 natürlich aufgrund der Art ihrer Berechnung demselben Verlauf folgen wie die durch eine SARS-CoV-2-Infektion verursachten und gemeldeten Todesfälle. Demnach gibt es in Deutschland zwischen Januar 2020 und April 2021 zwei Infektionswellen, die erste zwischen Februar und Mai 2020 mit bis zu über 15 000 Infektionen täglich und die zweite zwischen Oktober 2020 und März 2021 mit bis zu 65 000 täglichen Infektionen. Erste Infektionen tauchen Mitte Februar 2020 auf. Im Sommer 2020 gehen die Infektionszahlen auf bis zu knapp über 300 täglichen Infektionen zurück. Auch hier erkennt man den Beginn der dritten Welle ab Mitte März 2021.

Abschließend stellt die Abbildung (3.6) die auf wöchentlicher Basis aggregierten Infektionszahlen und die wöchentlich gemeldeten Infektionsfälle gegenüber. Die Daten mussten aggregiert werden, da die gegebenen Daten über die Positivrate der Tests und die Anzahl durchgeführter Tests, welche später für die Modellanalysen gebraucht werden, nur in wöchentlicher Form vorliegen. In dieser Abbildung und in allen weiteren Abbildungen dieser Arbeit, stellt jeweils das eingetragene Datum auf der x-Achse den letzten noch berücksichtigten Tag der Woche dar. Es wird deutlich, dass, auch wenn die blaue Kurve der gemeldeten Fälle demselben Verlauf folgt wie die rote Kurve der

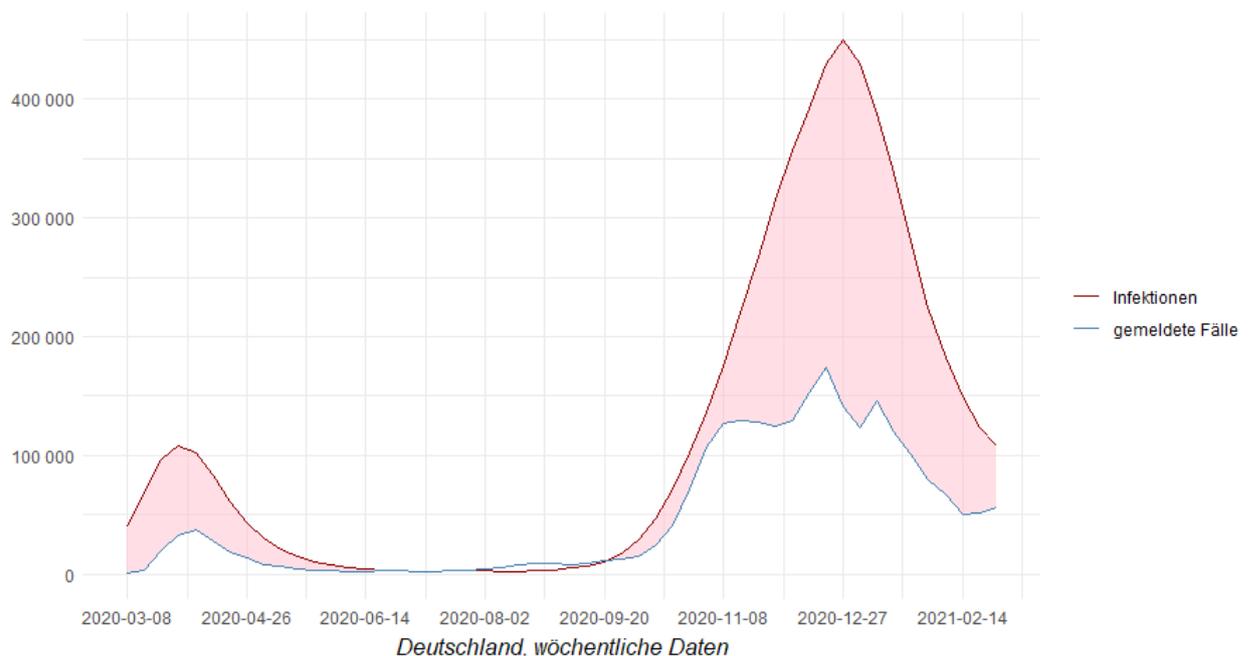


Abbildung 3.6: Infektionen vs. gemeldete Fälle in Deutschland

geschätzten tatsächlichen Infektionen, die gemeldeten Fälle überwiegend die tatsächlichen Infektionszahlen bei Weitem nicht abdecken. Während beider Infektionswellen werden weniger als die Hälfte aller Infektionen aufgedeckt. Nur in den Sommermonaten zwischen Juni und September 2020, wenn die Zahl der Infektionen sehr niedrig ist, liegen beide Kurven sehr nah beieinander. Teilweise beträgt die Zahl der gemeldeten Fälle höhere Werte als die Zahl der Neuinfektionen, was auf verschiedene Faktoren zurückgeführt werden kann: Ungenauigkeiten in der Berechnung der Infektionszahlen nach Abschnitt (2.1.3), Ungenauigkeiten bei der Meldung neuer Infektionsfälle sowie neuer Todesfälle, die Nichtberücksichtigung des *reporting delays* (siehe unteren Abschnitt (3.3)) und, wohl am wahrscheinlichsten, auf eine hohe Zahl falsch-positiver Testergebnisse (siehe Kapitel (4.2)). Diese Werte werden in dieser Arbeit gleich null gesetzt, das heißt es kann keine negative Dunkelziffer geben.

### 3.3 Reporting Delay

Vor allem im Bereich der Wochen vor dem 8. November 2020 ist der Abbildung (3.6) zu entnehmen, dass die Infektionszahlen noch um den *reporting delay* ergänzt werden sollten, um sie mit den gemeldeten Fallzahlen vergleichen zu können. Die Schätzungen des *reporting delays* sind in der Abbildung (3.7) visualisiert. Es wird deutlich, dass die meisten Fälle einen Meldeverzug von weniger als zehn Tagen aufweisen. Ca. 40% aller gemeldeten Fälle werden zwischen zwei und fünf Tagen gemeldet.

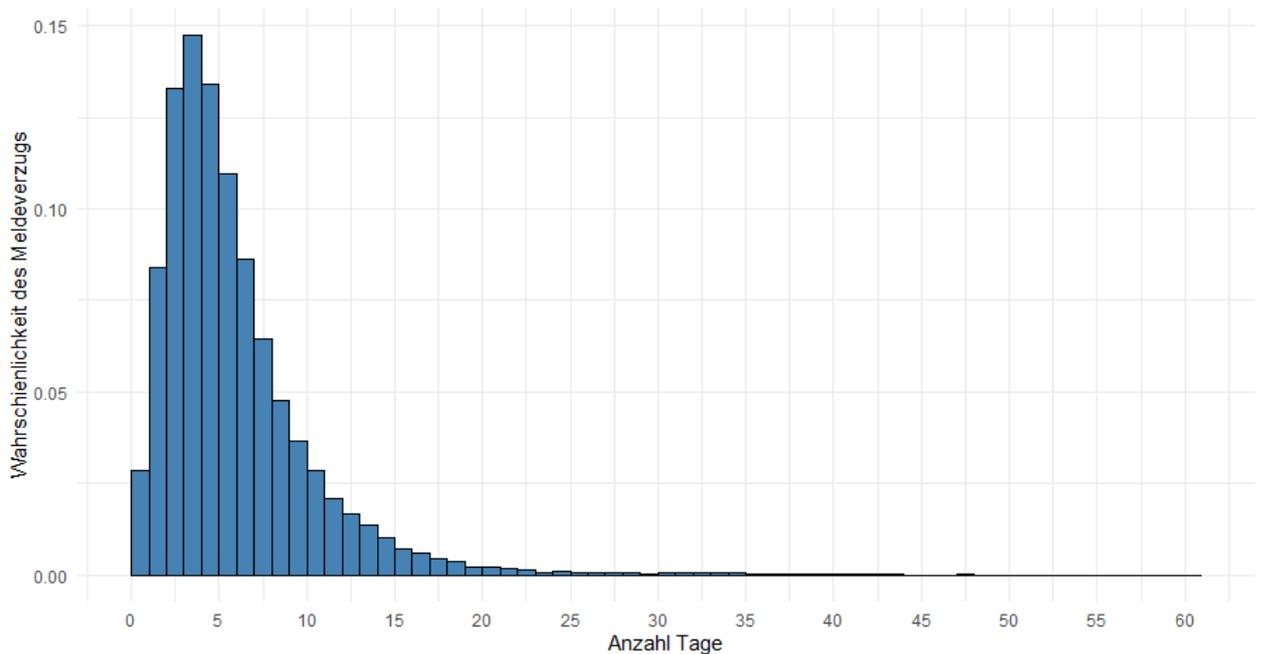


Abbildung 3.7: Schätzung der Verteilung der Zeit zwischen Symptombeginn und Meldung

Abbildung (3.8) zeigt das Ergebnis einer solchen Verschiebung der Infektionen um den *reporting delay*. Auch hier ist wieder im Bereich der Wochen vor dem 8. November 2020, als Infektionen und gemeldete Fallzahlen deutlich hochgehen, erkennbar, dass die Infektionen nun besser zu den gemeldeten Fallzahlen passen. Allerdings sieht man auch, dass die Verschiebung etwas zu groß war. Mithilfe dieser Steigung der Infektionen und gemeldeten Fallzahlen, kann durch Ausprobieren verschiedener Werte für den Meldeverzug der Todesfälle grafisch geschätzt werden, dass die Annahme einer Verzögerung der Meldung der Todesfälle von drei Tagen angemessen ist. Abbildung (3.9) zeigt das finale Ergebnis der Infektionszahlberechnung, in dem der *reporting delay* der Fallzahlen und der Meldeverzug der Todesfälle berücksichtigt worden sind.

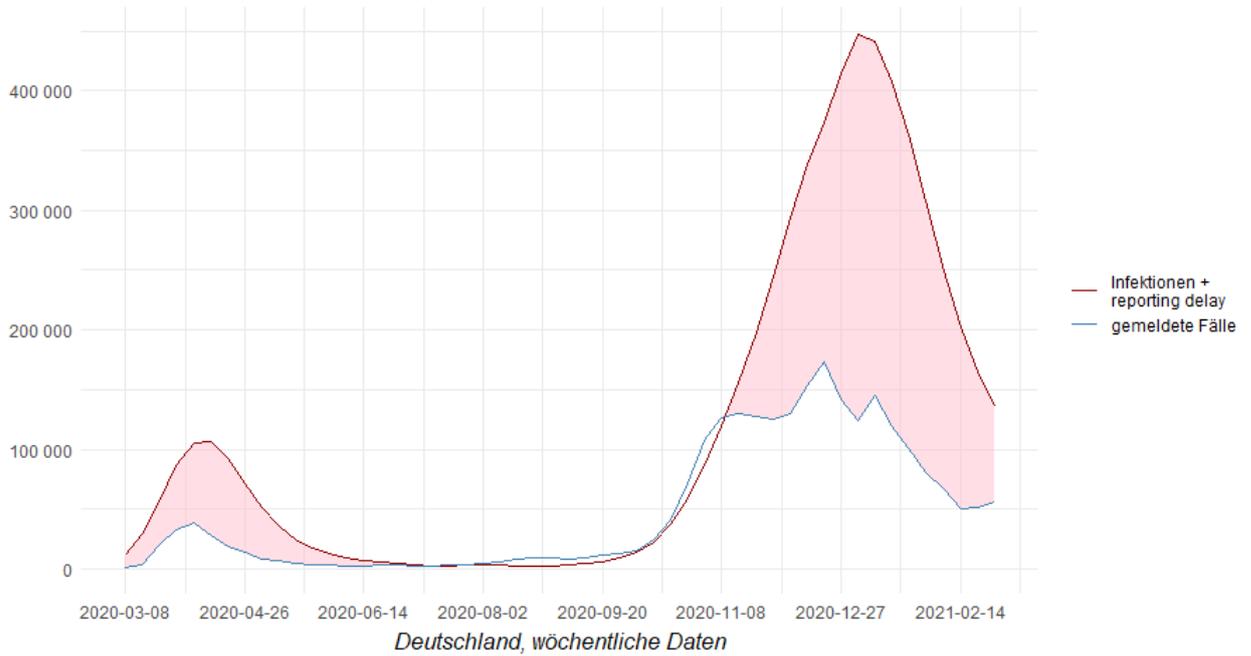


Abbildung 3.8: Modellierte Infektionen mit reporting delay der Infektionsfälle in Deutschland

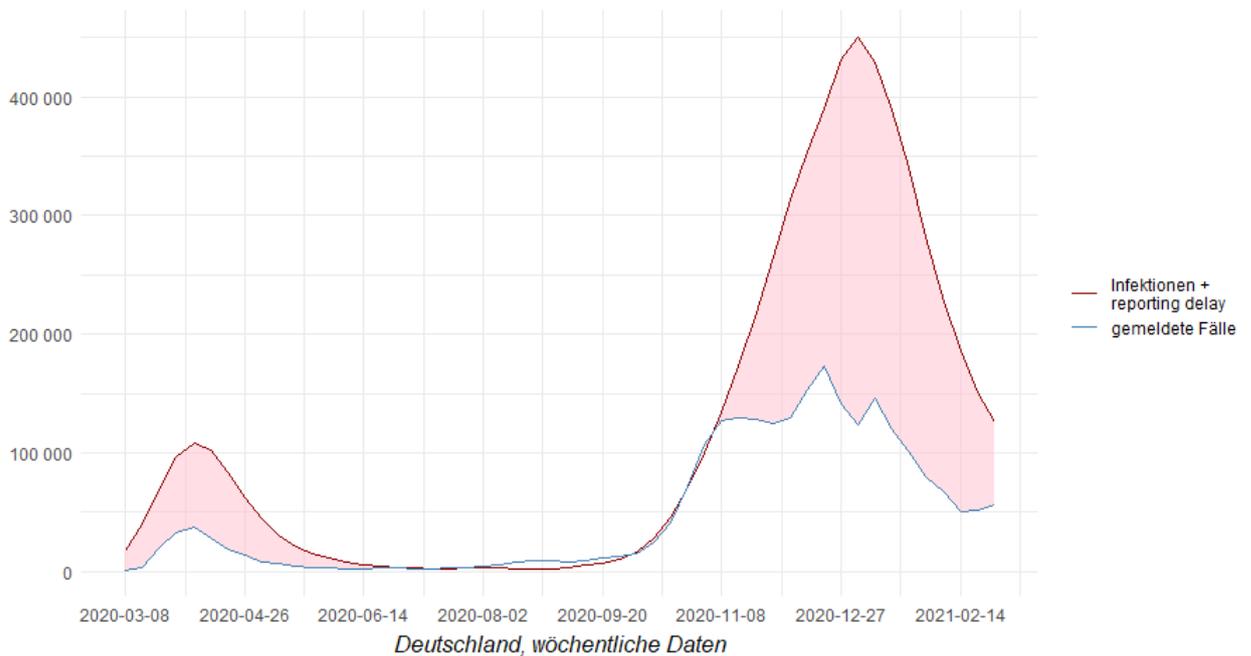


Abbildung 3.9: Modellierte Infektionen mit reporting delay der Infektionsfälle und der Todesfälle in Deutschland

## 3.4 Statische Modelle

### 3.4.1 Ein naives Modell

In einem ersten Versuch, den Einfluss der Testgröße und der Positivrate auf die Dunkelziffer in Deutschland zu analysieren, wird ein einfaches lineares Modell aufgestellt. Unsere Zielgröße ist dabei die Variable *Dunkelziffer*, welche die Zahl unentdeckter Infektionen abbildet, und unsere beiden Einflussgrößen sind die Variablen *Tests* und *Positivrate*, welche einmal die Zahl durchgeführter Tests pro 100000 Einwohner darstellt und einmal den Anteil positiver Tests an allen Tests in Prozent. Die Schätzungen der Parameter sind in der Tabelle (3.2) dargestellt.

$$\text{Dunkelziffer} = \alpha + \beta_1 \cdot \text{Tests} + \beta_2 \cdot \text{Positivrate} + \epsilon \quad (3.1)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	-13938.910	14695.940	-0.948	0.348
$\beta_1$	-11.070	13.960	-0.793	0.432
$\beta_2$	20709.180	1907.570	10.856	<1.23e-14 ***

signifikant auf einem Niveau von: \*\*\* = 0.001    \*\* = 0.01    \* = 0.05  
 $R^2 = 0.7408$ , adjustiertes  $R^2 = 0.7303$

Tabelle 3.2: Ergebnisse des naiven Modells

Dieses erste Modell besagt, dass die Positivrate einen positiven Einfluss auf die Dunkelziffer besitzt, während die Anzahl der durchgeführten Tests die Dunkelziffer negativ beeinflusst. Konkret besagt es, dass eine Erhöhung der Positivrate um einen Prozent, bei gleich bleibender Testgröße, in Erwartung eine Erhöhung der Dunkelziffer um rund 20709 unentdeckte Infektionen zur Folge hat. Zudem besagt es, dass bei einer Erhöhung der Testgröße um einen weiteren Test pro 100000 Einwohner und gleichzeitigem Festhalten der Positivrate, eine mittlere Verringerung der Dunkelziffer um 11 unentdeckte Infektionen zu erwarten ist. Allerdings ist die Parameterschätzung  $\beta_1$  für den Einfluss der Testgröße nicht signifikant, was bedeutet, dass die Nullhypothese  $\beta_1 = 0$  nicht abgelehnt werden kann. Somit ist bei diesem Modell nicht ausgeschlossen, dass die Testgröße überhaupt keinen Einfluss auf die Größe der Dunkelziffer ausübt. Dieses Modell weist mehrere Problematiken auf. Erstens ist womöglich die Zielgröße nicht ge-

eignet gewählt, da bspw. eine Dunkelziffer von 40 bei insgesamt 50 Neuinfektionen eine ganz andere Dynamik aufweist als die gleiche Dunkelziffer bei einer Gesamtzahl von bspw. 5000 Infektionen. Zudem erscheint es sehr unintuitiv, dass die Testgröße keinen Einfluss auf die Dunkelziffer hat. Um den Einfluss der Positivrate und der Testgröße auf die Dunkelziffer besser beurteilen zu können, werden nun beide Variablen getrennt voneinander betrachtet.

### 3.4.2 Lösungsansätze für den Einfluss der Testgröße

Möchte man nun den Einfluss der Testgröße einzeln auf die Dunkelziffer betrachten, ergibt sich folgende Problematik. Die Testgröße hat im zeitlichen Verlauf stark zugenommen, wie der linken Grafik aus Abbildung (3.10) zu entnehmen ist. Im gleichen

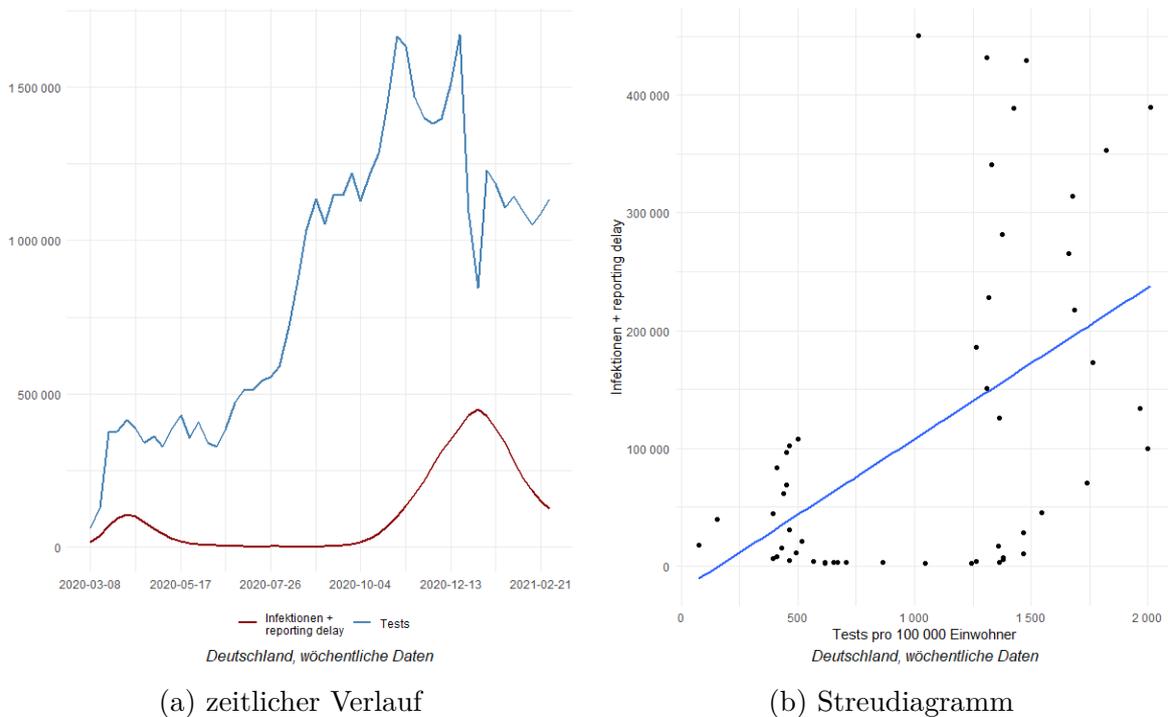


Abbildung 3.10: Infektionen vs. Anzahl Tests in Deutschland

Zeitraum hat aber auch die Zahl der Infektionen tendenziell zugenommen. Die zweite Welle im Winter zwischen 2020 und 2021 ist wesentlich stärker als die erste Anfang des Jahres 2020. Somit ergibt sich wie im Streudiagramm aus (3.10) abgelesen werden kann, ein positiver linearer Zusammenhang zwischen der Testgröße und der Zahl der Infektionen. Abbildung (3.11) zeigt eine noch wesentlich deutlichere Korrelation zwischen der Zahl der Neuinfektionen und der Dunkelziffer. Genauso wie die Zahl der

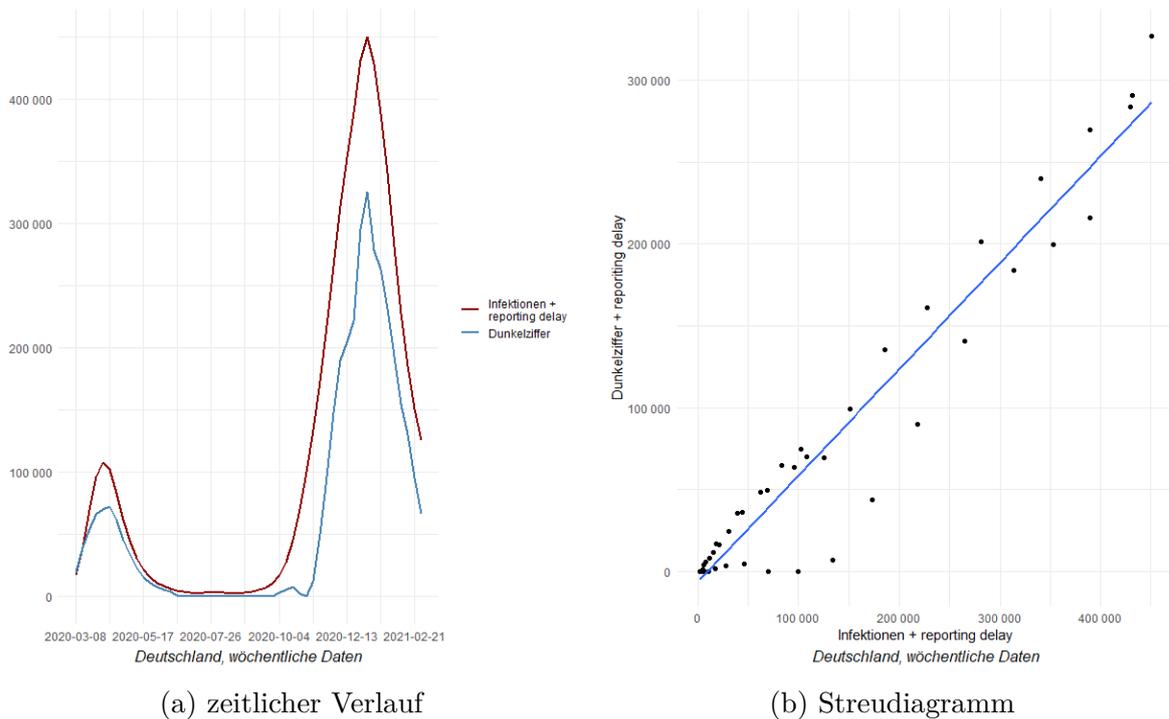


Abbildung 3.11: Infektionen vs. Dunkelziffer in Deutschland

Neuinfektionen verläuft die Dunkelziffer in zwei Wellen, wobei die zweite deutlich größer ausfällt als die erste. Sieht man sich rechts davon das dazugehörige Streudiagramm an, wird dieser Zusammenhang noch deutlicher. Ergo ergibt sich auch ein positiver Zusammenhang zwischen der Testgröße und der Dunkelziffer. Dieser lässt sich auch mit Abbildung (3.12) darstellen.

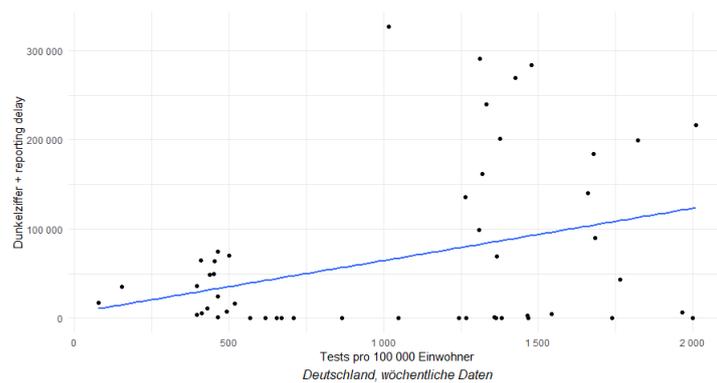


Abbildung 3.12: Anzahl Tests vs. Dunkelziffer in Deutschland

Folglich würde eine lineare Spezifikation à la  $Dunkelziffer = \alpha + \beta_1 \cdot Tests + \epsilon$  der Testgröße einen positiven Koeffizienten zuweisen, was natürlich unlogisch ist. Der Grund,

warum die Testgröße und die Dunkelziffer positiv miteinander korrelieren, liegt darin, dass beide mit der Zahl der Infektionen zusammenhängen. Die Zahl der Neuinfektionen beeinflusst stark die Zahl neuer nicht erkannter Fälle. Die Pandemie hat im Verlauf der Zeit an Dynamik dazugewonnen, gleichzeitig sind die Testkapazitäten in Deutschland immer weiter ausgebaut worden und immer mehr Tests sind durchgeführt worden. Dies führt zur positiven Korrelation zwischen der Zahl der Infektionen und der Zahl der durchgeführten Tests. Folglich muss davon ausgegangen werden, dass eine Erhöhung der Zahl durchgeführter Tests die Dunkelziffer nicht erhöht, es handelt sich hier nur um Korrelation, nicht um Kausalität. Im Folgenden werden zwei Lösungsversuche zur Adressierung dieses Problems vorgestellt, bei denen sich herausstellt, dass der zweite Ansatz zu bevorzugen ist.

### Lösungsansatz 1: Anteil der Dunkelziffer an Infektionen

Um den Effekt der Infektionszahlen auf die Dunkelziffer zu entfernen, wird nicht mehr die Dunkelziffer selber, sondern ihr Anteil an den Infektionszahlen als Zielvariable betrachtet. Abbildung (3.13) zeigt ein Streudiagramm, welches auf der x-Achse die Zahl

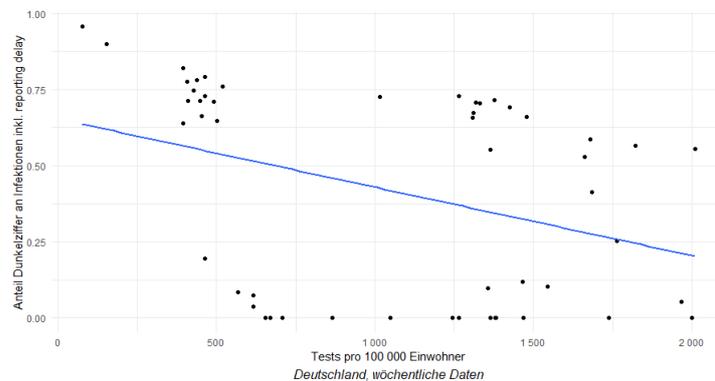


Abbildung 3.13: Anteil Dunkelziffer vs. Infektionen in Deutschland

der durchgeführten Tests pro 100000 Einwohner und auf der y-Achse den relativen Anteil der Dunkelziffer an den Infektionszahlen, also  $\frac{\text{Dunkelziffer}}{\text{Infektionen}}$ , abbildet. Zwar zeigt sich jetzt ein negativer Zusammenhang zwischen Testgröße und relativer Dunkelziffer, allerdings erkennt man klare Strukturen im zeitlichen Verlauf, je nachdem in welchem Zeitpunkt der Pandemie wir uns befinden, ob in der ersten Welle, der zweiten Welle, oder den ruhigen Sommermonaten dazwischen. Es bleibt zu bedenken, dass genau wie eine Dunkelziffer von 40 bei 50 Infektionen anders zu bewerten ist als bei 50000 Infektionen, auch die Betrachtung der absoluten Testgröße die gleichen Schwächen aufweist.

40 durchgeführte Tests bei 50 Neuinfektionen sind nicht zu vergleichen mit einer Testgröße von 40 durchgeführten Tests bei 50000 Neuinfektionen. Diese Überlegung führt uns nun zum zweiten Lösungsansatz.

### Lösungsansatz 2: Anteil Tests und Anteil Dunkelziffer an Infektionen

In einem zweiten Ansatz wird nun nicht mehr nur die Dunkelziffer relativ an der Zahl der Neuinfektionen gemessen, sondern auch die Zahl der durchgeführten Tests wird in Relation zu der Zahl vorhandener Neuinfektionen gesetzt. Die linke Grafik in Abbil-

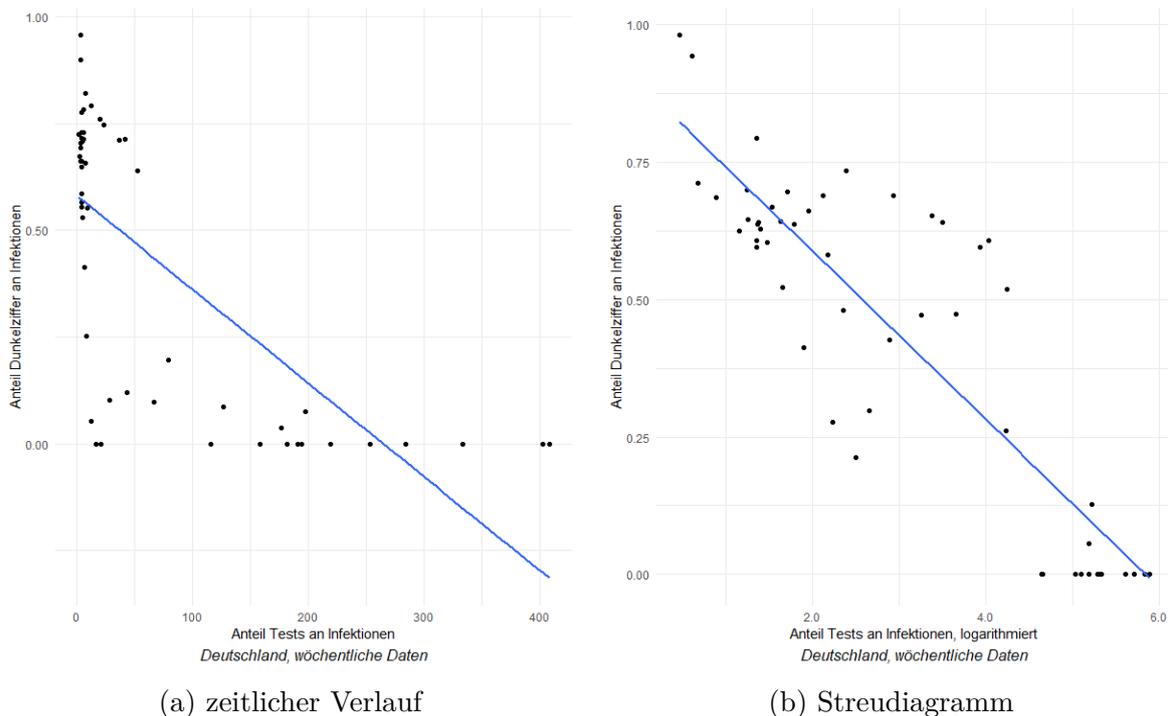


Abbildung 3.14: Anteil Tests und Anteil Dunkelziffer in Deutschland

dung (3.14) zeigt ein Streudiagramm der beiden Größen *Anteil Tests an Infektionen* und *Anteil Dunkelziffer an Infektionen*. Diese wurden berechnet, indem jeweils die Zahl der Dunkelziffer und die Gesamtzahl durchgeführter Tests durch die Zahl der Neuinfektionen geteilt wurden. Es zeigt sich, dass die Testgröße bis zu 300 Mal größer ist als die Zahl der Neuinfektionen, dies ist den Sommermonaten des Jahres 2020 zu verschulden. Währenddessen nimmt der Anteil der Dunkelziffer selbstverständlich nur Werte zwischen 0 und 1 an. Deutlich wird ein negativer Zusammenhang zwischen dem Anteil der Tests und der Dunkelziffer sichtbar, ab einem Faktor von über 100 um den die Testgröße größer ist als die Zahl der Infektionen, gibt es kaum noch unentdeckte Fäl-

le. Dieser Zusammenhang scheint allerdings nicht linear zu sein. Da die Größe *Anteil Tests* linkssteil verteilt zu sein scheint, wird diese Größe im nächsten Schritt mit dem natürlichen Logarithmus, der in dieser Arbeit mit  $\log()$  gekennzeichnet wird, logarithmiert. Das Ergebnis sieht man in der rechten Grafik von Abbildung (3.14). Diese zeigt ein Streudiagramm zwischen der logarithmierten relativen Testgröße und der relativen Dunkelziffer. Ein lineares Modell zum Modellieren dieses Zusammenhangs ist nun nach dem Logarithmieren der Testgröße besser geeignet. Im Großen und Ganzen lässt sich also sagen, dass dieser Lösungsansatz, das Betrachten der relativen Dunkelziffer und der logarithmierten relativen Testgröße zu bevorzugen ist. Im nächsten Abschnitt wird dieses Modell angewandt und die zugehörigen Koeffizienten geschätzt.

### Modell für den Einfluss der Testgröße

$$\text{Anteil der Dunkelziffer} = \alpha + \beta_1 \cdot \log(\text{Anteil der Tests}) + \epsilon \quad (3.2)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	0.90390	0.05591	16.167	<2e-16 ***
$\beta_1$	-0.16092	0.01621	-9.925	2.07e-13 ***

signifikant auf einem Niveau von: \*\*\* = 0.001   \*\* = 0.01   \* = 0.05  
 $R^2 = 0.6633$ , adjustiertes  $R^2 = 0.6566$

Tabelle 3.3: Ergebnisse des Testgrößenmodells

Dieses Modell besagt, dass der logarithmierte relative Anteil der Testgröße den relativen Anteil der Dunkelziffer negativ beeinflusst. Das bedeutet, dass eine Erhöhung der Testgröße relativ zur Anzahl der Neuinfektionen den relativen Anteil der Dunkelziffer an den Infektionen verringert. Die genaue Interpretation des geschätzten Koeffizienten  $\beta_1$  ist nicht auf den ersten Blick erkennbar. Eine Erhöhung der logarithmierten relativen Testgröße führt zu einer erwarteten Verringerung des Anteils der Dunkelziffer an den Infektionen um rund 0.16. Um diesen Zusammenhang besser interpretieren zu können, werden folgenden Überlegungen aufgestellt:

$$\begin{aligned} \log(a \cdot b) &= \log(a) + \log(b) \\ \log(\exp^1) &= 1 \end{aligned}$$

$$\Rightarrow \log(a \cdot \exp^1) = \log(a) + 1$$

Da  $\exp^1 \approx 2.72$ , kann die Koeffizientenschätzung aus Tabelle (3.3) folgendermaßen interpretiert werden. Erhöht sich die relative Testgröße um den Faktor 2.72, so verringert sich der relative Anteil der Dunkelziffer im Mittel um 0.16. Das bedeutet schließlich, dass Veränderungen der relativen Testgröße, wenn diese eher klein ist, zu größeren Veränderungen bei der relativen Dunkelziffer führen. Eine Erhöhung des Anteils der Testgröße von 2 auf rund 5 ( $2 \cdot 2.72 \approx 5$ ), führt zu einer Verringerung der relativen Dunkelziffer um 0.16. Gleichzeitig führt auch eine Erhöhung der relativen Testgröße von 200 auf 500 zur genau gleichen Verringerung der relativen Dunkelziffer. Abbildung (3.15) zeigt mit der blauen Kurve den tatsächlichen Anteil der Dunkelziffer an den Infektionen, während die rote Kurve den mit dem Testgrößenmodell geschätzten Anteil der Dunkelziffer beschreibt.

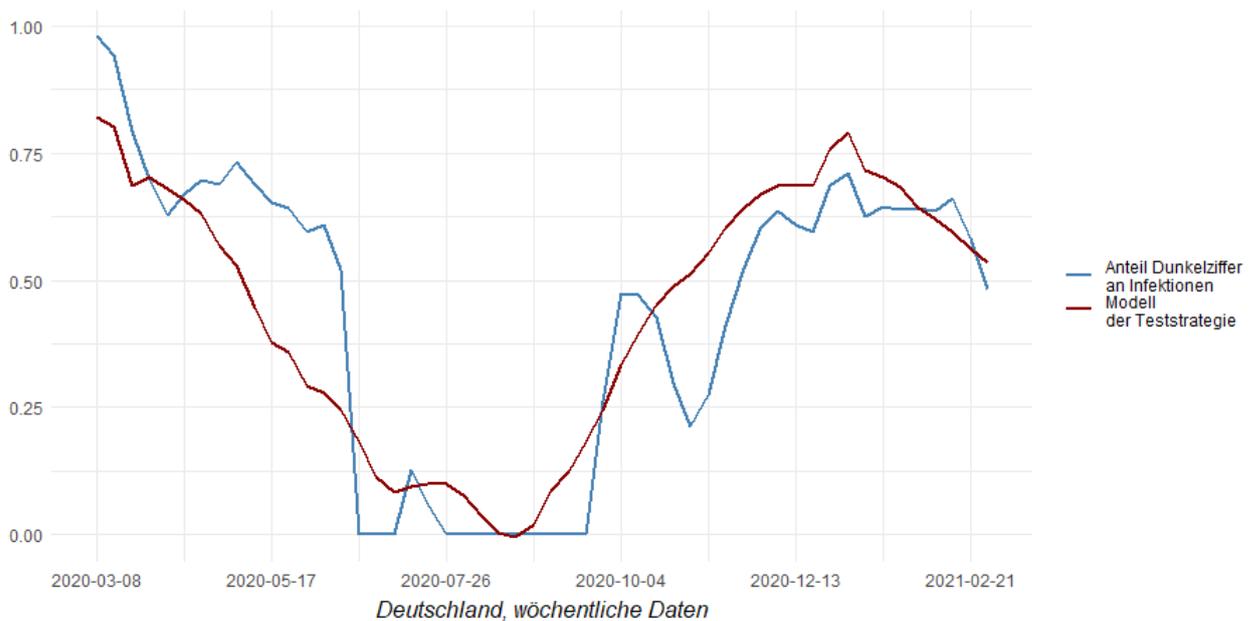


Abbildung 3.15: Modellschätzung durch Testgröße in Deutschland

**Alternative Definition der Dunkelziffer**

$$\log\left(\frac{\text{nicht entdeckte Fälle}}{\text{entdeckte Fälle}} + 1\right) = \alpha + \beta_1 \cdot \log\left(\frac{\text{Tetsts}}{\text{Infektionen}}\right) + \epsilon \quad (3.3)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	1.66913	0.13997	11.925	3.12e-16 ***
$\beta_1$	-0.30425	0.04059	-7.495	1.01e-09 ***

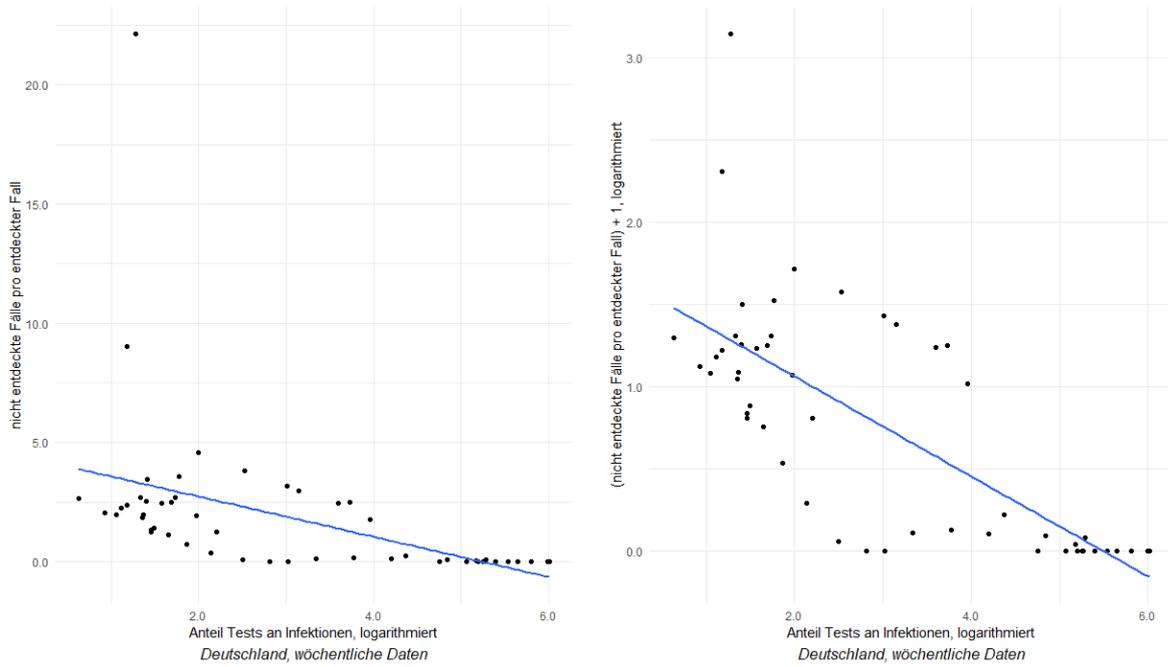
signifikant auf einem Niveau von: \*\*\* = 0.001    \*\* = 0.01    \* = 0.05  
 $R^2 = 0.5291$ , adjustiertes  $R^2 = 0.5197$

Tabelle 3.4: Ergebnisse des Testgrößenmodells mit alternativer Definition der Dunkelziffer

Zur schöneren Interpretation der Variablen, kann die Dunkelziffer als Quotient:  $\frac{\text{nicht entdeckte Fälle}}{\text{entdeckte Fälle}}$  definiert werden. Somit gibt sie an, wie viele unentdeckte Fälle es pro entdeckten Fall gibt. Der Anteil der Tests bleibt so definiert wie bisher. Die Streudiagramme in (3.16) zeigen, dass wieder weitere Variablentransformationen nötig sind, um einen linearen Zusammenhang herzustellen. Zusätzlich zum Logarithmieren der Einflussvariable *Anteil der Tests*, ist dieses Mal auch ein Logarithmieren der neuen Zielvariable notwendig. Vor dem Logarithmieren muss zusätzlich eine 1 auf die Relation  $\frac{\text{nicht entdeckte Fälle}}{\text{entdeckte Fälle}}$  dazuaddiert werden, da diese Relation auch Werte von 0 annimmt. Man erhält die untere Modellgleichung. Das Logarithmieren der Zielgröße verbessert die Modellschätzung erheblich, das  $R^2$  steigt dabei von 0.1869 auf 0.5291, auch die Signifikanz der Einflussvariable steigt deutlich.

Die Ergebnisse aus (3.4) zeigen, dass, auch durch die neue Definition der Dunkelziffer, der Einfluss der Testgröße auf die Dunkelziffer signifikant bleibt. Die genaue Interpretation der Koeffizientenschätzer gestaltet sich schwierig. Analog zu vorherigem Modell gilt: eine Erhöhung der relativen Testgröße um der Faktor 2,72, verringert die um 1 erhöhte und logarithmierte Zahl der unentdeckten Fälle pro entdecktem Fall im Mittel um rund 0,3. An dieser Stelle wird deutlich, dass die neue Definition der Dunkelziffer nicht zu einer schöneren Interpretation des Modells geführt hat, ganz im Gegenteil, sie ist deutlich erschwert worden. Dies und die durch das Modell verminderte erklärte Streuung (erkennbar durch die Reduktion des  $R^2$ ), zeigen die Unterlegenheit dieses

Modells zum vorherigen auf. Deswegen wird es nicht weiter verwendet in dieser Arbeit.



(a) Dunkelziffer nicht logarithmiert

(b) Dunkelziffer logarithmiert

Abbildung 3.16: Streudiagramme der alternativen Definition der Dunkelziffer

### 3.4.3 Lösungsansätze für den Einfluss der Positivrate

#### Erster Lösungsansatz

Auch bei der Betrachtung des Einflusses der Positivrate auf die Dunkelziffer haben sich zwei mögliche Modelle ergeben. Im ersten Modell (3.4) wird ein linearer Zusammenhang zwischen der Positivrate in % und der absoluten Dunkelziffer angenommen.

$$\text{Dunkelziffer} = \alpha + \beta_1 \cdot \text{Positivrate} + \epsilon \quad (3.4)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	-22378	10100	-2.216	0.0313 *
$\beta_1$	20015	1689	11.853	< 3.9e-16 ***

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.7375$ , adjustiertes  $R^2 = 0.7323$

Tabelle 3.5: Ergebnisse des Positivrate-Dunkelziffer-Modells

Die Ergebnisse aus Tabelle (3.5) zeigen, dass ein solches Modell zu sehr signifikanten Koeffizientenschätzungen führt, auch das Modell selbst kann mit einem  $R^2$  von über 0.7 also mehr als 70% der Streuung erklären. Dieses Modell sagt also die Dunkelziffer überraschend gut voraus. Es besagt, dass eine Erhöhung der Positivrate um einen Prozent, die Dunkelziffer im Mittel um 20015 unentdeckte Fälle erhöht. Denkt man über diesen Zusammenhang nach, und ruft sich nochmal ins Gedächtnis, dass die Dunkelziffer in Deutschland sehr stark mit den Neuinfektion korreliert (siehe Abbildung (3.11)), liegt es nahe, dass die Zahl der Infektionen in Deutschland die Positivrate der durchgeführten Tests direkt stark beeinflusst, was natürlich inhaltlich Sinn ergibt. Deshalb wird im nächsten Abschnitt noch ein zweites Modell vorgestellt, welches den Zusammenhang der Positivrate direkt mit der Zahl der Neuinfektionen modelliert.

#### Zweiter Lösungsansatz

Abbildung (3.17) zeigt, dass ein starker Zusammenhang zwischen den Infektionen und der Positivrate der Tests vorliegt. Im Umkehrschluss kann man dann davon ausgehen, dass mithilfe der Positivrate die Zahl der Infektionen, die in dieser Arbeit mithilfe der Todesfälle zurückgerechnet worden sind, geschätzt werden können.

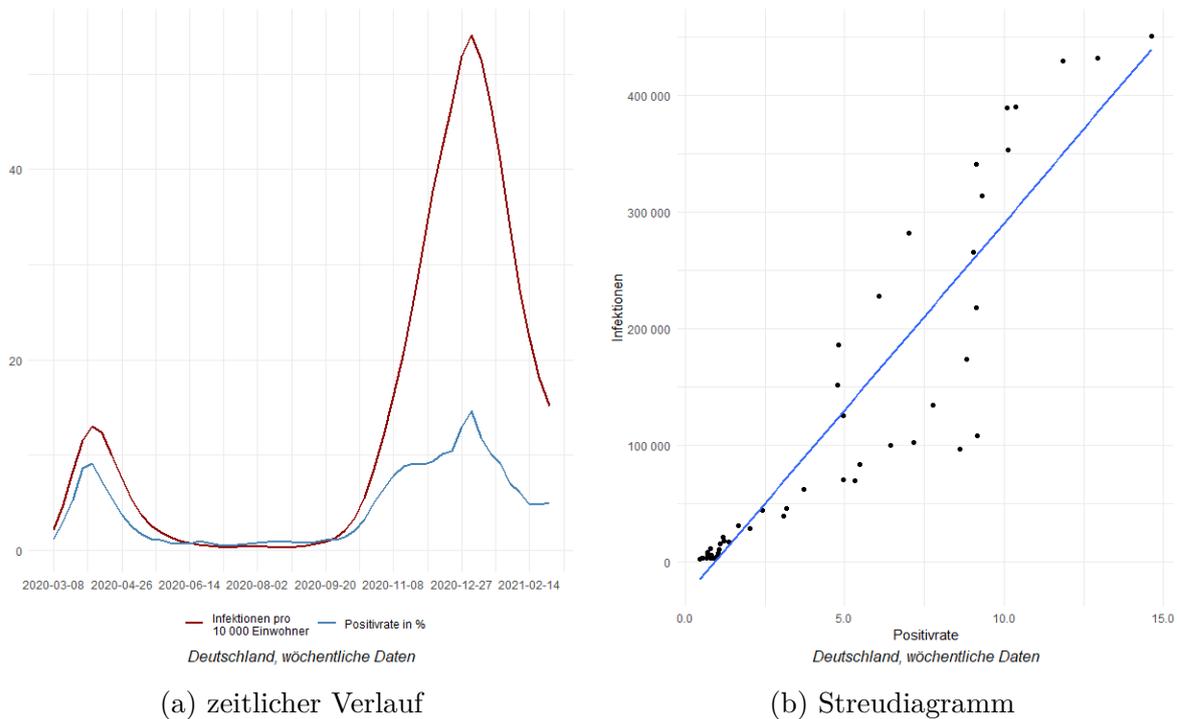


Abbildung 3.17: Infektionen vs. Positivrate in Deutschland

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \epsilon \tag{3.5}$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	-30269	11394	-2.657	0.0106 *
$\beta_1$	32074	1905	16.837	< 2e-16 ***

signifikant auf einem Niveau von: \*\*\* = 0.001    \*\* = 0.01    \* = 0.05  
 $R^2 = 0.8501$ , adjustiertes  $R^2 = 0.8471$

Tabelle 3.6: Ergebnisse des Positivrate-Infektionen-Modells

Die Ergebnisse dieses Positivrate-Infektionen-Modells sind in der Tabelle (3.6) zu entnehmen. Auch hier ist das Ergebnis für den Koeffizientenschätzer  $\beta_1$  hoch signifikant. Das auf fast 0.9 erhöhte  $R^2$  legt nahe, dass die Positivrate primär von der Zahl der Infektionen beeinflusst wird. Dieses Modell sagt aus, dass bei einer Erhöhung der Positivrate um einen Prozent, die Zahl der Infektionen im Mittel um 32074 steigt.

**Vergleich beider Modelle der Positivrate**

Die Abbildungen (3.18) und (3.19) zeigen einmal die Schätzung der Dunkelziffer nur mithilfe der Positivrate und einmal die Schätzung der Infektionszahlen nur mithilfe der Positivrate. Es zeigt sich, dass beide Modelle zu sehr guten Schätzungen führen, auch wenn das Modell der Infektionszahlen besser passt. Sowohl die Dunkelziffer als auch die Infektionszahlen werden zur Zeit der ersten Welle überschätzt. Betrachtet man vorherige Abbildungen, die den Verlauf der Positivrate, der Dunkelziffer und der Infektionen darstellen ((3.11), (3.1)), so erkennt man, dass die Positivrate der Tests während der ersten Welle, zur Zeit hoher Infektionen aber sehr weniger Tests, nicht so gravierend niedriger ist als die Positivrate während der zweiten Welle, als es noch sehr viel höhere Infektionszahlen, unentdeckte Fälle und auch Tests gibt. Das heißt durch diese Rahmenbedingungen (weniger Tests, generell weniger Infektionszahlen) scheint sich die Positivrate in der ersten Welle anders auf die Zahl der Infektionen auszuwirken als während der zweiten Welle. Ob das an gezielterem Testen aufgrund weniger Testmöglichkeiten liegt oder andere Gründe hat, bleibt spekulativ. In jedem Fall scheint das Positivrate-Infektionen-Modell gegenüber dem Positivrate-Dunkelziffer-Modell überlegen zu sein und wird daher bevorzugt. Es ist interessant, wie gut die Zahl der Infektionen nur anhand der Positivrate geschätzt werden kann, ohne die Testgröße und somit die gemeldeten Fallzahlen zu berücksichtigen.

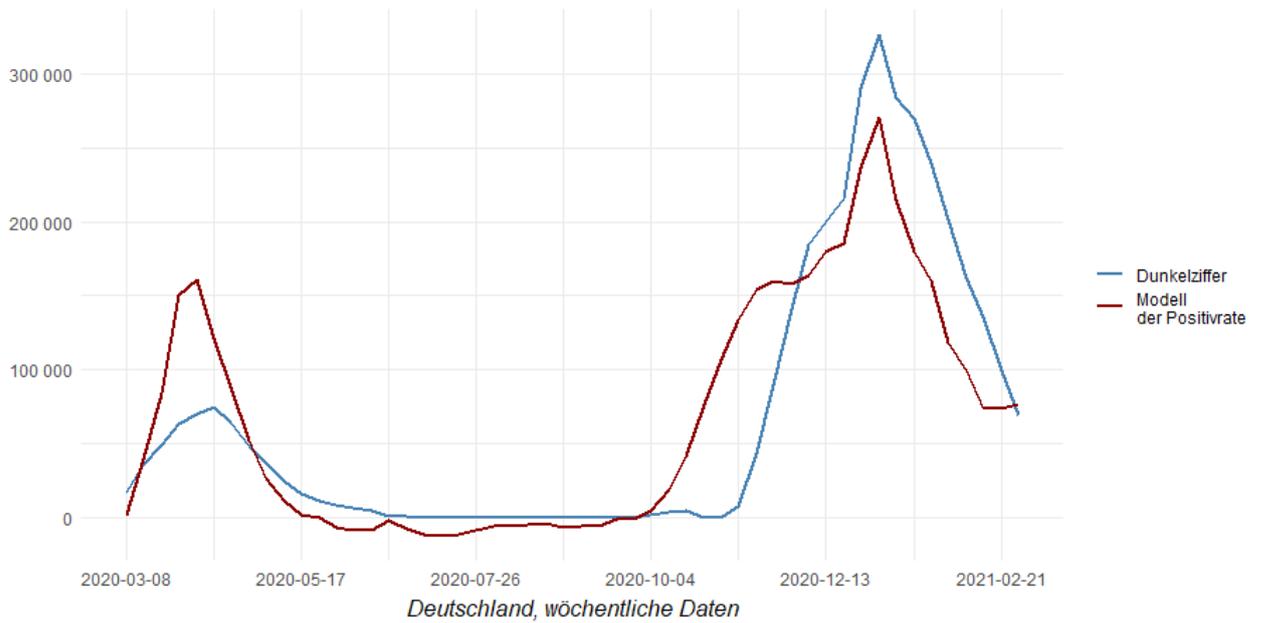


Abbildung 3.18: Modellschätzung der Dunkelziffer durch Positivrate in Deutschland

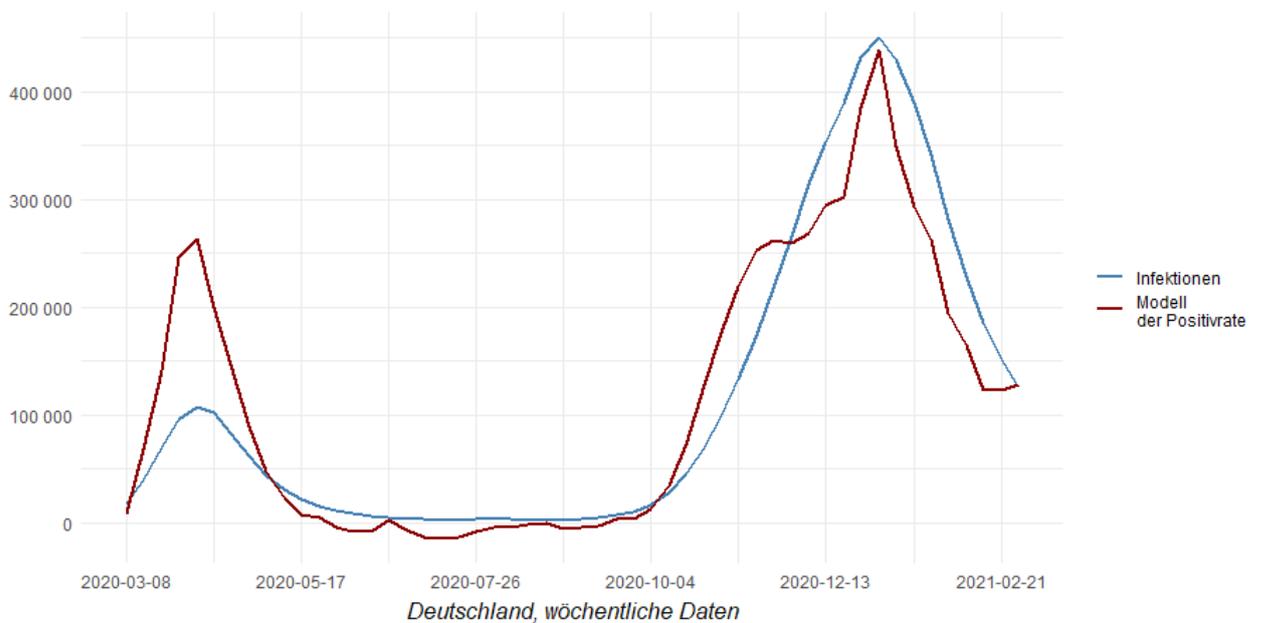


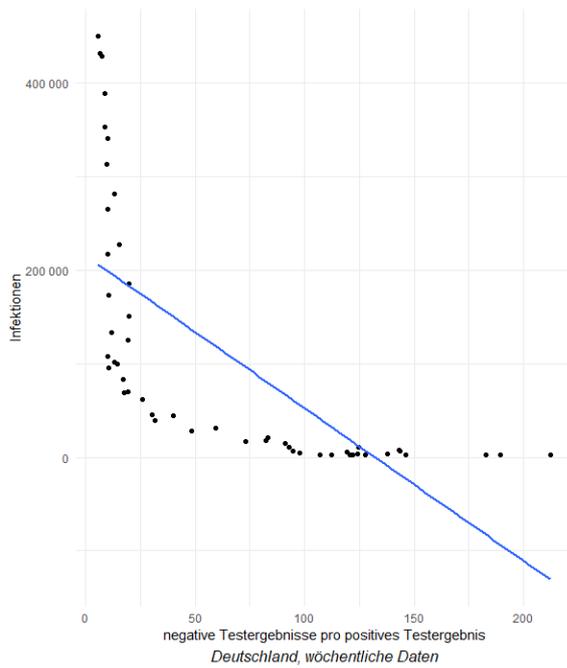
Abbildung 3.19: Modellschätzung der Infektionen durch Positivrate in Deutschland

### Alternative Definition der Positivrate

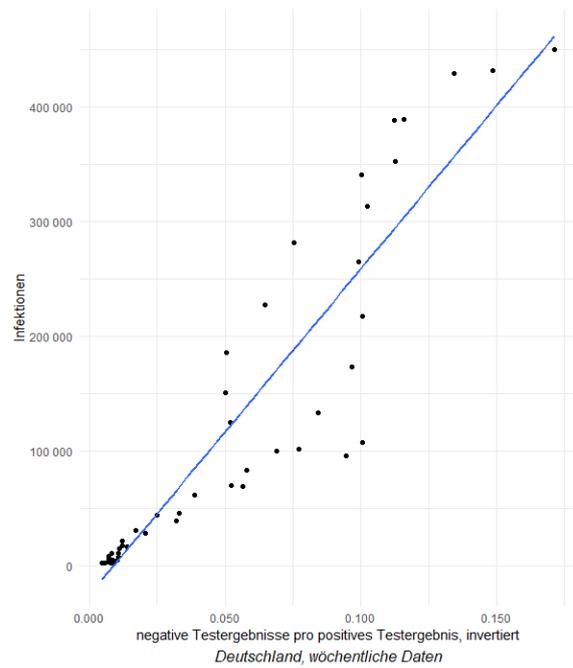
Analog zu Abschnitt (3.4.2) kann man sich überlegen, dass es auch schönere Definitionen für die Positivrate gibt. Anstatt dem Quotienten  $\frac{\text{gemeldete Fälle}}{\text{Testgrö.,e}}$  zu betrachten, kann man auch die Relation von negativen Testergebnissen zu positiven Testergebnissen betrachten oder auch die Relation von durchgeführten Tests zu positiven Tests. Der Vorteil hierbei ist, dass diese neue Arten der Definition greifbarer sind. Bei einer, nach bisheriger Definition, Testpositivrate von 10%, wäre die Relation von negativen Testergebnissen zu positiven Testergebnissen bei einem Wert von neun. Das heißt pro positives Testergebnis gibt es neun negative. Die Relation von durchgeführten Tests zu positiven Tests läge bei zehn, das heißt es müssen im Mittel zehn Tests durchgeführt werden, um ein positives Ergebnis zu erhalten.

Allerdings zeigen die Streudiagramme in Abbildung (3.20), dass bei einer solchen Definition der Positivrate die Positivrate invertiert werden sollte, um einen optimalen linearen Zusammenhang zwischen der Positivrate und der Zahl der Infektionen herzustellen. Man hätte dann bei einer Inversion beider Alternativen der Positivrate wieder nur Werte unter eins, da die Zahl positiver Tests die Zahl negativer Tests immer übersteigt und natürlich auch die Zahl der durchgeführten Tests immer größer ist als die Zahl positiver Ergebnisse. Bei der zweiten Alternative hat man natürlich nach einer Inversion sowieso wieder unsere ursprüngliche Definition der Positivrate. Dies führt dazu, dass auch im *Positivrate-Infektionen-Modell* die ursprüngliche Definition der Positivrate beibehalten wird.

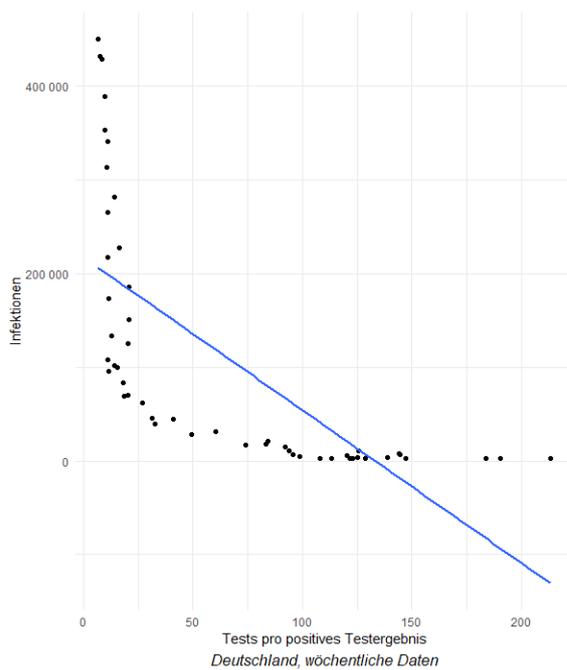
Bevor im übernächsten Kapitel weitere Modelle zur Analyse des Einflusses von Testgröße und Testpositivrate auf die Dunkelziffer unter Berücksichtigung zeitlich verzögerter Einflüsse vorgestellt werden (sogenannte *dynamische Modelle*), werden nun im nächsten Kapitel Methoden zur Evaluierung von Zeitreihenmodellen vorgestellt und auf die in diesem Kapitel vorgestellten *statischen Modelle* angewandt.



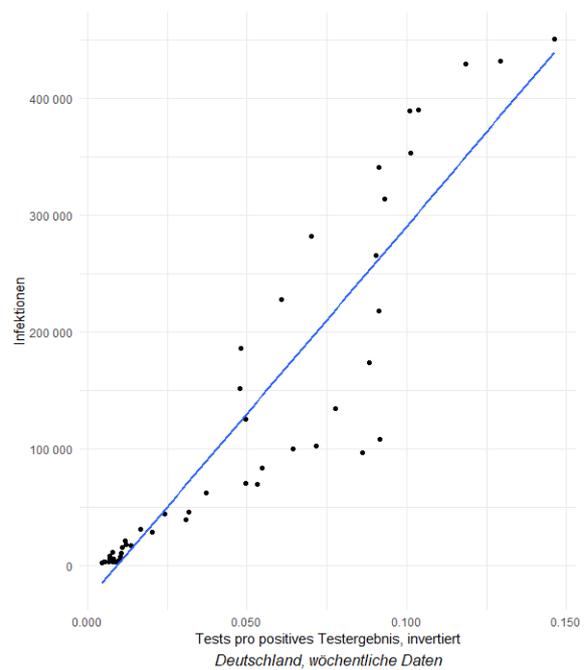
(a) Residuenanalyse des naiven Modells



(b) Residuenanalyse des Testgrößen - Modells



(c) Residuenanalyse des Positivrate - Dunkelziffer - Modells



(d) Residuenanalyse des Positivrate - Infektionen - Modells

Abbildung 3.20: Streudiagramme der alternativen Definitionen der Positivrate und der Infektionen

## 3.5 Evaluationen der statischen Modelle

### 3.5.1 Teil I: Autokorrelation der Fehler

In der Abbildung (3.21) sind vier Grafiken zu allen vier im Kapitel (3.4) angewandten Modellen abgebildet. Jede der vier Grafiken enthält wiederum drei Grafiken. Die obere bildet die Residuen des Modells im zeitlichen Verlauf ab. Die x-Achse ist dabei die Zeitachse in der alle betrachteten Wochen abgetragen sind. Die linke untere Abbildung stellt den im Methoden- und Theorieteil beschriebenen ACF-Plot dar und die rechte visualisiert die Häufigkeitsverteilung der Residuen. Die Abbildung der Residuen gegen die Zeit weist in allen Modellen bis auf das Testgrößenmodell zwei Ausreißer nach unten auf, und zwar nach der fünften und nach der fünfunddreißigsten Woche nach Beobachtungsbeginn. Hier werden also die Dunkelziffer bzw. die Infektionen durch die Modelle überschätzt. Diese Zeitpunkte stellen einmal den Höhepunkt der Infektionen in der ersten Welle dar und einmal den Zeitraum des Beginns der zweiten Welle (vgl. dazu die Abbildungen (3.18, 3.19)). Dafür überschätzen diese drei Modelle die Dunkelziffer bzw. Infektionen im Bereich der zweiten Welle. Die Residuen des Testgrößenmodells sind dagegen zufälliger verteilt. Im Zeitraum der ersten Welle unterschätzt dieses Modell eher, im Gegensatz zu den anderen drei Modellen, den Anteil der Dunkelziffer. Die Normalverteilungsannahme der Residuen ist nach der Grafik unten links bei allen Modellen im Großen und Ganzen gegeben. Anhand der ACF-Plots ist zu erkennen, dass alle Modelle eine Autokorrelation der Residuen aufweisen. Das Testgrößenmodell weist diese sogar bis zur sechsten Ordnung, während die restlichen drei Modelle diese bis zur dritten Ordnung aufweisen. Auch der Breusch-Godfrey-Test gibt für das *naive Modell* einen p-Wert von  $7.197 \cdot 10^{-7}$ , für das Testgrößenmodell einen von  $2.663 \cdot 10^{-6}$ , für das Positivrate-Deunkelziffer-Modell einen von  $7.029 \cdot 10^{-7}$  und für das Positivrate-Infektionen-Modell eine von  $2.603 \cdot 10^{-6}$  aus. Folglich kann für jedes Modell die Nullhypothese abgelehnt werden, dass keine Autokorrelation der Fehler bis zur zehnten Ordnung vorliegt. Das bedeutet, dass die Schätzungen der Standardabweichungen der Koeffizientenschätzungen aus Kapitel (3.4) fehlerhaft sind. Um dieses Problem der Autokorrelation der Fehler zu beheben, werden im Kapitel (3.6) die statischen Modelle durch Hinzufügen von *Lags* spezifiziert. Hier wird jedoch zunächst eine andere Methode des Umgangs mit Autokorrelation der Fehler angewandt, nämlich die Neuschätzung der Standardabweichungen mit der *Newey-West-Methode*. Die Ergebnisse der *Newey-West-Methode* sind in den Tabellen (3.7) und (3.9) aufgeführt. Diese neuen Schätzer für die Varianzen der geschätzten Koeffizienten sind auch robust

Koeffizient	$\sqrt{\text{Var}}$	t-Wert	P(> t )
$\alpha$	0.08089	11.174	3.362e-15 ***
$\beta_1$	0.01347	-11.947	2.906e-16 ***

Tabelle 3.7: Newey-West-Schätzung, Testgrößenmodell

Koeffizient	$\sqrt{\text{Var}}$	t-Wert	P(> t )
$\alpha$	0.05591	16.167	<2e-16 ***
$\beta_1$	0.01621	-9.925	2.07e-13 ***

Tabelle 3.8: gewöhnliche Schätzung, Testgrößenmodell

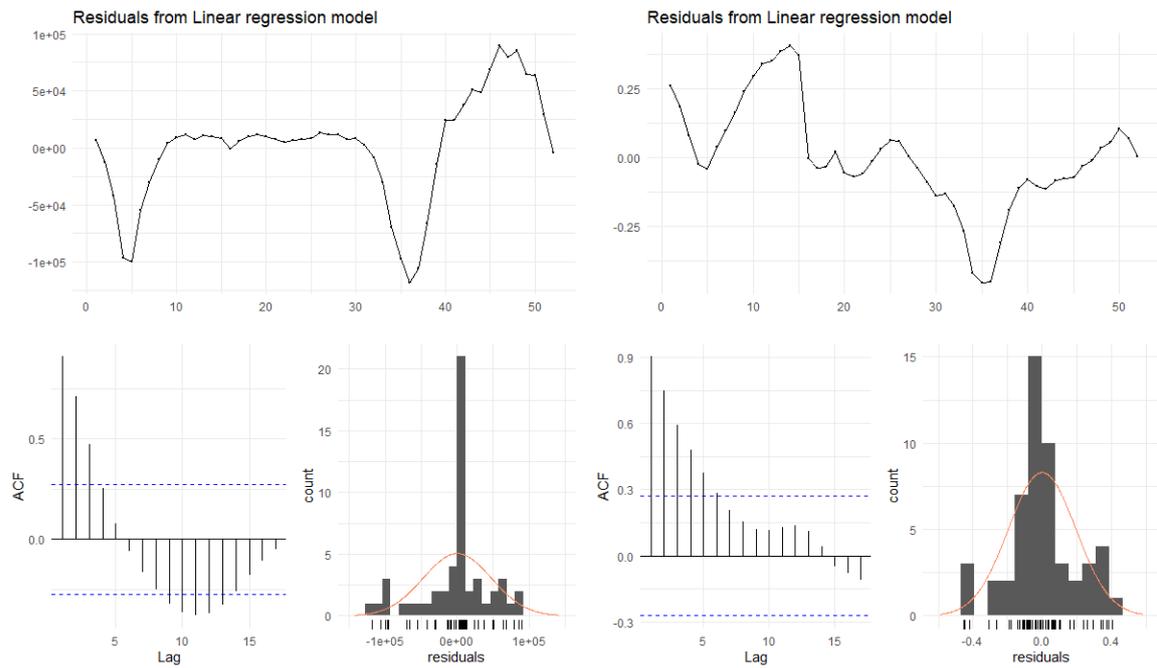
Koeffizient	$\sqrt{\text{Var}}$	t-Wert	P(> t )
$\alpha$	7102.1	-4.2619	8.958e-05 ***
$\beta_1$	3435.5	9.3361	1.540e-12 ***

Tabelle 3.9: Newey-West-Schätzung, Positivrate-Infektionen-Modell

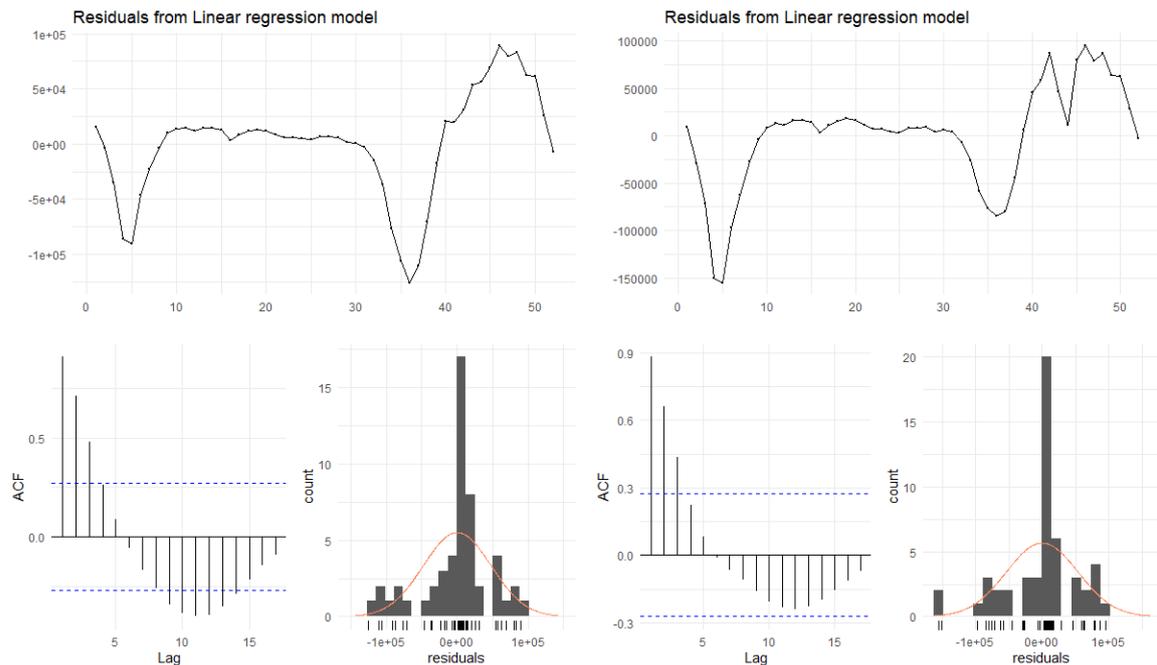
Koeffizient	$\sqrt{\text{Var}}$	t-Wert	P(> t )
$\alpha$	11394	-2.657	0.0106 *
$\beta_1$	1905	16.837	< 2e-16 ***

Tabelle 3.10: gewöhnliche Schätzung, Positivrate-Infektionen-Modell

gegen Heteroskedastie, welche bei allen statischen Modellen, wie der Abbildung (3.22) zu entnehmen ist, vorkommt. Verwendet man für die Schätzung von  $L$  die in der Literatur empfohlene Formel  $L = 0.75 \cdot n^{\frac{1}{3}}$ , erhalte man  $L \approx 3$ . Unser  $n$  ist dabei 52, da wir für Deutschland 52 Wochen betrachten. Die Abbildung (3.21) zeigt jedoch, dass für das Testgrößenmodell eine Autokorrelation der Residuen bis zur sechsten Ordnung vermutet werden kann, deshalb wir für dieses Modell ein  $L$  von 6 benutzt. Die unteren Tabellen zeigen, dass, sowohl für das Testgrößenmodell als auch für das Positivrate-Infektionenmodell, die neue Schätzung zur leichten Veränderung der Varianz der Koeffizienten, deren t-Werte und Signifikanzen führt. Dabei verringert sich die Varianz der Koeffizientenschätzung der Einflussvariable  $\log(\text{Anteil der Testgröße})$  für das Testgrößenmodell, was zu einer leicht erhöhten Signifikanz dieses Koeffizientenschätzers führt. Bei dem Positivrate-Infektionenmodell hingegen erhöht sich die Varianz der Koeffizientenschätzung der Einflussvariable *Positivrate*, was zu einer leicht verringerten Signifikanz führt. Im Großen und Ganzen ändern sich die Schätzungen mit der neuen Methode nicht gravierend, und beide Schätzungen des Effekts der Einflussgrößen bleiben hoch signifikant auf einem Niveau von 0,1 %.

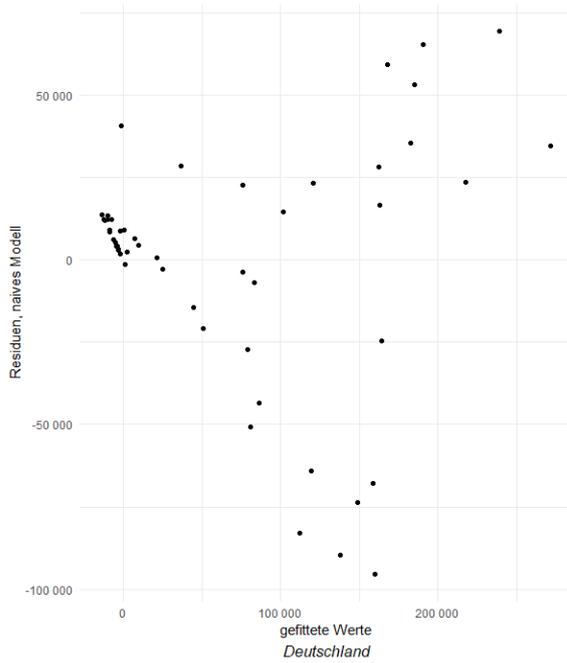


(a) Residuenanalyse des naiven Modells (b) Residuenanalyse des Testgrößen - Modells

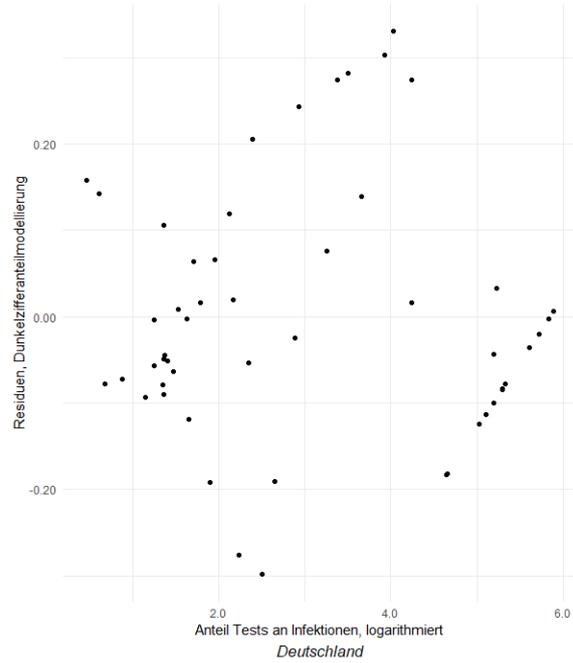


(c) Residuenanalyse des Positivrate - Dunkelziffer - Modells (d) Residuenanalyse des Positivrate - Infektionen - Modells

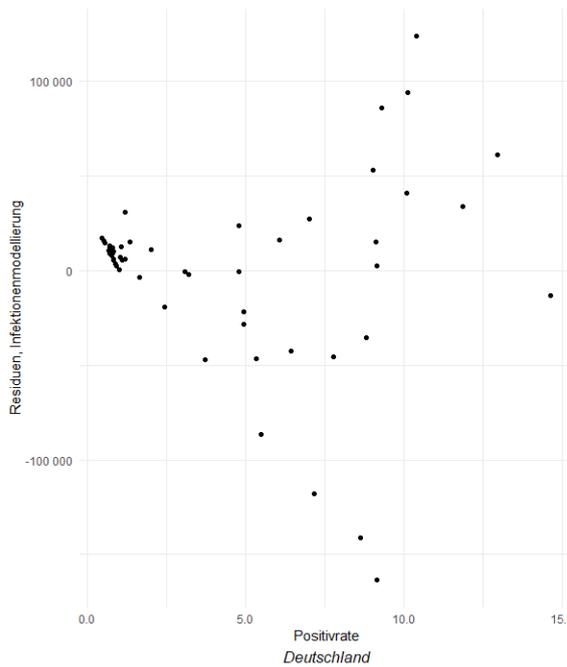
Abbildung 3.21: Residualanalyse der einfachen statischen Modelle  
jeweils oben: Plot der Residuen gegen die Zeit, jeweils unten links: ACF-Plot, jeweils unten rechts: Häufigkeitsverteilung der Residuen



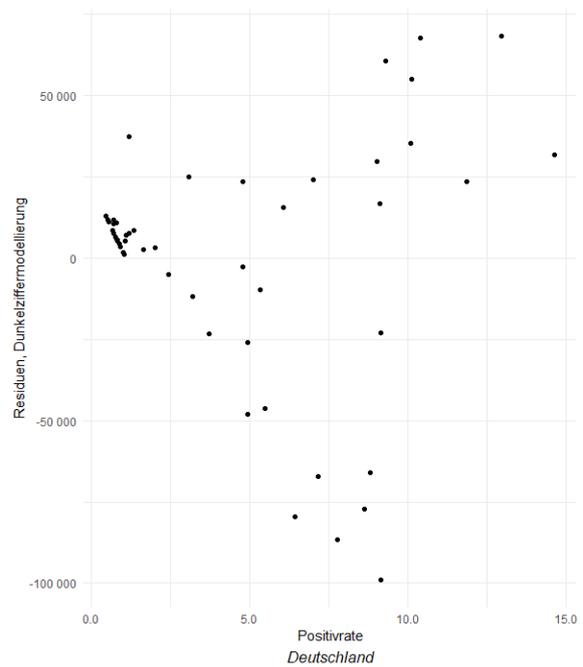
(a) naives Modell, gefittete Werte



(b) Testgrößen-Modell



(c) Positivrate-Infektionen-Modell



(d) Positivrate-Dunkelziffer-Modell

Abbildung 3.22: Residuenplots der statischen Modelle

### 3.5.2 Teil II: Stationarität der Variablen

Betrachtet man den Verlauf unserer Variablen *Infektionen*, *Dunkelziffer*, *Testgröße* und *Positivrate* wird klar, dass sie alle nicht stationär sind. Während der ersten und zweiten Infektionswelle steigen sowohl Infektionszahlen, als auch die Dunkelziffer sowie die Positivrate. Die Testgröße nimmt im Verlauf des betrachteten Zeitraums stetig zu. Eine Möglichkeit zur Herstellung von Stationarität stellt die Differenzbildung nicht-stationärer Variablen dar.

$$\text{Diff}(\text{Anteil der Dunkelziffer}) = \alpha + \beta_1 \cdot \text{Diff}(\log(\text{Anteil der Tests})) + \epsilon \quad (3.6)$$

$$\text{Diff}(\text{Infektionen}) = \alpha + \beta_1 \cdot \text{Diff}(\text{Positivrate}) + \epsilon \quad (3.7)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\hat{\alpha}$	-0.005086	0.011173	-0.455	0.650950
$\hat{\beta}_1$	-0.157134	0.037870	-4.149	0.000132 ***

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.26$ , adjustiertes  $R^2 = 0.2449$

Tabelle 3.11: Ergebnisse des Differenzen-Testgrößenmodells

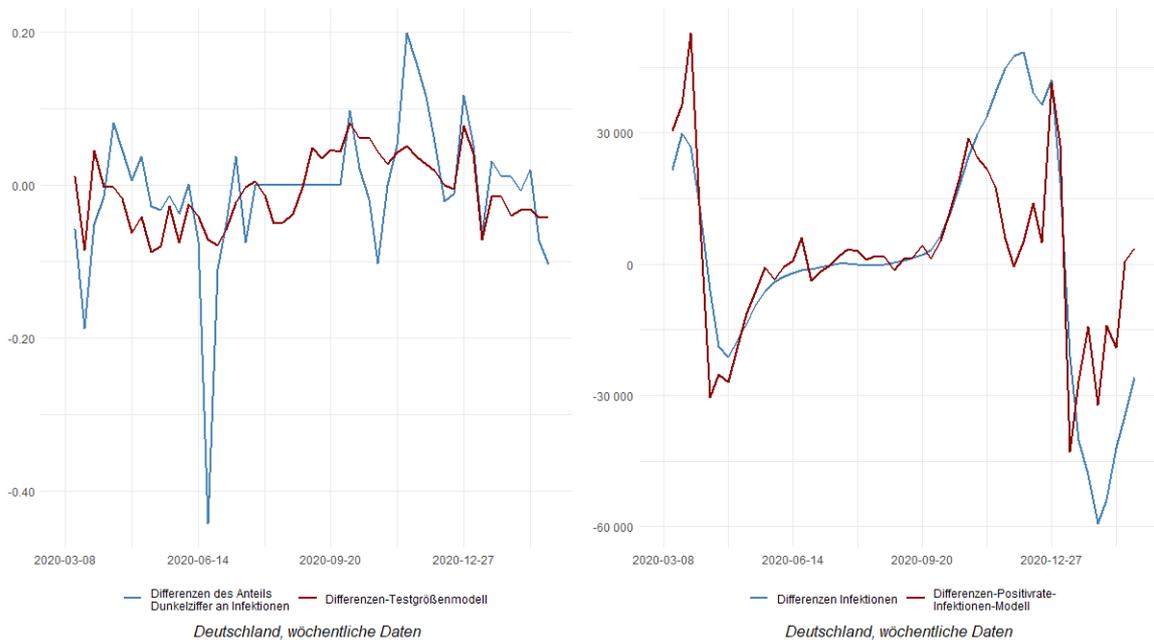
Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\hat{\alpha}$	953.4	2571.9	0.371	0.712
$\hat{\beta}_1$	15746.3	2170.0	7.256	2.66e-09 ***

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.518$ , adjustiertes  $R^2 = 0.5081$

Tabelle 3.12: Ergebnisse des Differenzen-Positivrate-Infektionen-Modells

Die Ergebnisse beider Modellgleichungen zeigen sehr signifikante Ergebnisse für den Einfluss der Differenz der logarithmierten relativen Testgröße sowie der Differenz der Positivrate. Das Differenzen-Testgrößenmodell besagt, dass bei einer Erhöhung der logarithmierten Differenz der relativen Testgröße zwischen zwei aufeinanderfolgenden Wochen um den Faktor 2,72, sich die Differenz der relativen Dunkelziffer zwischen den gleichen Wochen im Mittel um ca. 0,16 verringert. Das Differenzen-Positivrate-Infektionen-Modell besagt, dass bei einer Erhöhung der Differenz der Positivrate zweier

aufeinanderfolgender Wochen um einen Prozent, die Differenz der Infektionen zwischen den beiden Wochen im Mittel um ca. 15746 steigt. In beiden Modellen ist dabei die Differenz zwischen zwei Wochen so zu verstehen, dass von einer Woche der Wert der vorangegangenen Woche abgezogen wird. Im Vergleich zu den Modellen ohne Differenzen schneiden diese beiden Modelle jedoch deutlich schlechter ab. Das  $R^2$  sinkt deutlich von 0.6633 auf 0,26 bzw. von 0,8501 auf 0,518. Die Schätzung durch die beiden Modelle wird in Abbildung (3.23) veranschaulicht.

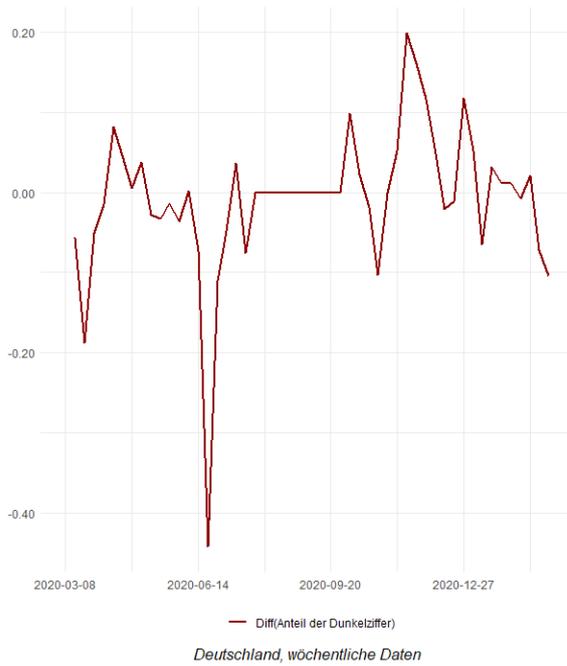


(a) Differenz relative Dunkelziffer

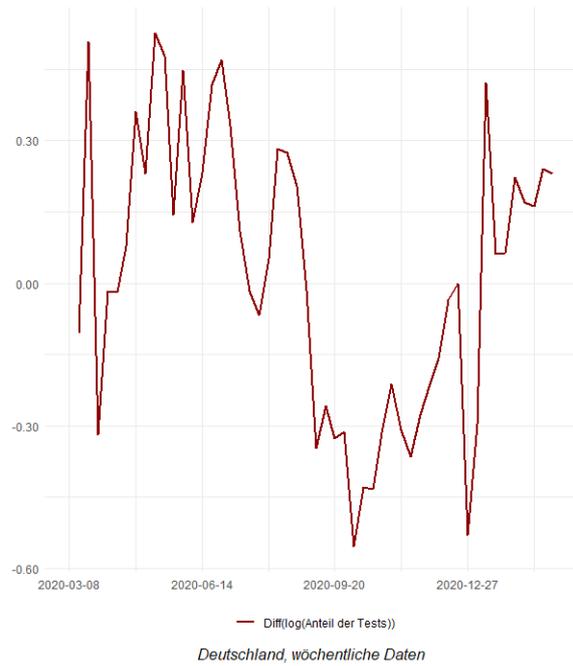
(b) Differenz logarithmierte relative Testgröße

Abbildung 3.23: Schätzungen der Differenzenmodelle

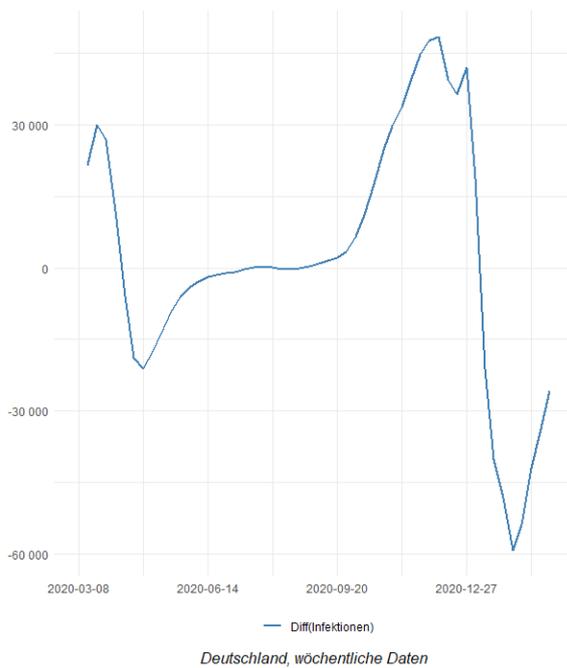
Die Grafiken der Abbildung (3.24) zeigen, dass bei allen Variablen, außer bei den Infektionszahlen, durch Bildung erster Differenzen die Nicht-Stationarität eliminiert bzw. zumindest stark reduziert werden konnte. Bei den Infektionszahlen könnte man auch zweite Differenzen bilden, also nochmal Differenzen der Differenzen selber. Dies erscheint jedoch nicht so sinnvoll, da das Modell noch schwerer zu interpretieren wird und allein schon durch die Bildung erster Differenzen sich die Modellgüte stark verschlechtert hat. Trotzdem kann aus den beiden Modellen mitgenommen werden, dass trotz der überwiegenden Eliminierung der Trends unserer Variablen, der Einfluss von Positivrate und Testgröße signifikant geblieben ist.



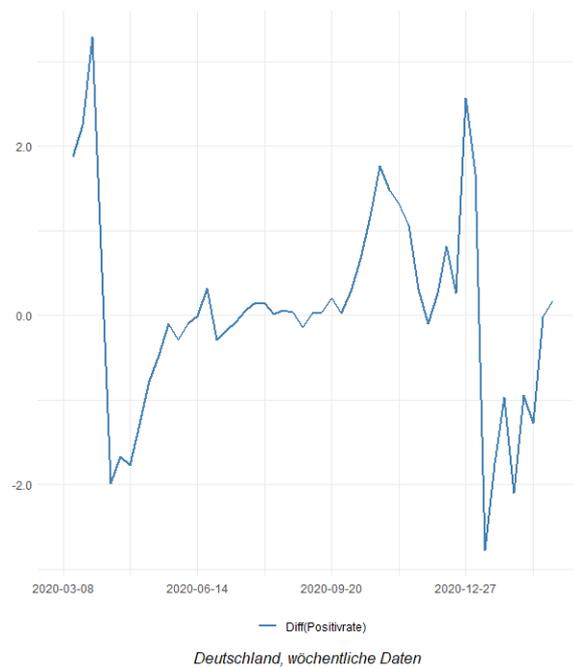
(a) Differenz relative Dunkelziffer



(b) Differenz logarithmierte relative Testgröße



(c) Differenz Infektionen



(d) Differenz Positivrate

Abbildung 3.24: Bildung erster Differenzen

Das Problem der Nicht-Stationarität unserer trendbehafteten Variablen besteht darin, dass anzunehmen ist, dass diese Variablen aus inhaltlichen Überlegungen einen gemeinsamen Trend besitzen. Es ist aus logischen Überlegungen heraus auszuschließen, dass der Trend in den Infektionszahlen und in der Positivrate nicht zusammenhängen, nur von einer Scheinkorrelation auszugehen, wäre nicht logisch. Daher erscheint das sonst übliche Vorgehen der Differenzenbildung zur Eliminierung von Nicht-Stationarität in unserem Fall nicht so sinnvoll. Erste Differenzen sind kurzfristige Veränderungen, langfristige Zusammenhänge werden dabei eliminiert. Genau an diesen langfristigen Zusammenhängen sind wir jedoch interessiert, wenn wir verstehen wollen, wie Infektionszahlen und Positivrate sich beeinflussen oder wie die Testgröße und die Dunkelziffer zusammenhängen. Andererseits ist deutlich zu erkennen, dass vor allem das Positivrate-Infektionen-Modell abhängig vom Zeitraum, den Zusammenhang zwischen Infektionszahlen und Positivrate unterschiedlich gut modelliert. Im Zeitraum der ersten Infektionswelle Anfang 2020 werden die Infektionszahlen deutlich überschätzt, während das Modell für den Zeitraum der zweiten Infektionswellen Ende 2020 und Anfang 2021 besser funktioniert. Grund dafür ist eben die fehlende Stationarität wichtiger Variablen, welche die Positivrate der Tests und die Infektionszahlen beeinflussen. Die Zahl der Infektionen und die Zahl der durchgeführten Tests unterscheiden sich grundlegend zwischen diesen beiden Zeiträumen. Auch unterscheiden sich diese Variablen wiederum grundlegend in den Sommermonaten dazwischen.

Deswegen wird im nächsten Schritt ein neues Modell unter der Berücksichtigung von Strukturbrüchen aufgestellt. Unser Positivrate-Infektionen-Modell wird um zwei weitere Variablen *erste Welle* und *zweite Welle* erweitert, welche nur den Wert 0 oder 1 annehmen können. Die Variable *erste Welle* nimmt dabei den Wert 1, wenn die Datenpunkte sich im Zeitraum der ersten Welle befinden, ansonsten ist sie 0. Die Variable *zweite Welle* hingegen nimmt dann den Wert 1 an, wenn die Datenpunkte sich im Zeitraum der zweiten Welle befinden, ansonsten ist sie 0. Ein Setzen beider Variablen auf 0 beschreibt die Sommermonate zwischen den beiden Wellen.

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \beta_2 \cdot \text{erste Welle} + \beta_3 \cdot \text{zweite Welle} + \epsilon \quad (3.8)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	-18866	11943	-1.580	0.1207
$\beta_1$	30432	2470	12.320	<2e-16 ***
$\beta_2$	-50733	20173	-2.515	0.0153 *
$\beta_3$	17319	22559	0.768	0.4464

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.8852$ , adjustiertes  $R^2 = 0.878$

Tabelle 3.13: Ergebnisse des Positivrate-Infektionen-Modells mit Strukturbrüchen

Diese Modellschätzung liefert signifikante Ergebnisse für den Effekt der *Positivrate* als auch der *ersten Welle*. Auch hat sich das  $R^2$  im Vergleich zum gleichen Modell ohne die Einbeziehung von Strukturbrüchen leicht von 0.8501 auf 0.8852 erhöht. Abbildung (3.25) zeigt jedoch, dass nur die Einführung eines Intercepts für die drei verschiedenen Zeitabschnitte nicht ausreicht. Es ist zu erkennen, dass trotz einer Niveaushiftung der Schätzung der Infektionen durch die Positivrate nach unten, das Modell für die *erste Welle* nicht so gut funktioniert wie für die restlichen Monate. Es ist naheliegend, dass die Art, wie die Positivrate die Zahl der Infektionen beeinflusst, während der ersten Welle eine andere ist als z.B. während der zweiten Welle. Deswegen wird nun ein neues Modell aufgestellt, welches Interaktionseffekte zwischen zeitlichen Abschnitten und der Positivrate berücksichtigt.

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \beta_2 \cdot \text{erste Welle} + \beta_3 \cdot \text{zweite Welle} + \beta_4 \cdot \text{Positivrate} \cdot \text{erste Welle} + \epsilon \quad (3.9)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	-23603	9355	-2.523	0.0151 *
$\beta_1$	36613	2217	16.517	< 2e-16 ***
$\beta_2$	33945	21745	1.561	0.1252
$\beta_3$	-23577	19034	-1.239	0.2216
$\beta_4$	-25323	4487	-5.644	9.27e-07 ***

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.9316$ , adjustiertes  $R^2 = 0.9258$

Tabelle 3.14: Ergebnisse des Positivrate-Infektionen-Modells mit Strukturbrüchen

Dieses Modell schätzt die Zahl der Infektionen nochmal deutlich besser als die vorgegangen. Auch die Abbildung (3.26) zeigt, dass durch dieses Modell systematische zeitliche Verzerrungen eliminiert werden konnten. Das Modell liefert signifikante Ergebnisse für den Einfluss der Positivrate sowie des Interaktionseffektes aus Positivrate und erster Welle. Es besagt, dass in den Sommermonaten und während der zweiten Welle eine Erhöhung der Positivrate um einen Prozent im Mittel eine Erhöhung der Infektionszahlen um 36613 zur Folge hat. Während der ersten Welle hat eine Erhöhung der Positivrate um einen Prozent im Mittel bloß eine Erhöhung der Infektionszahlen um (36613 – 25323), also um 11290 zur Folge.

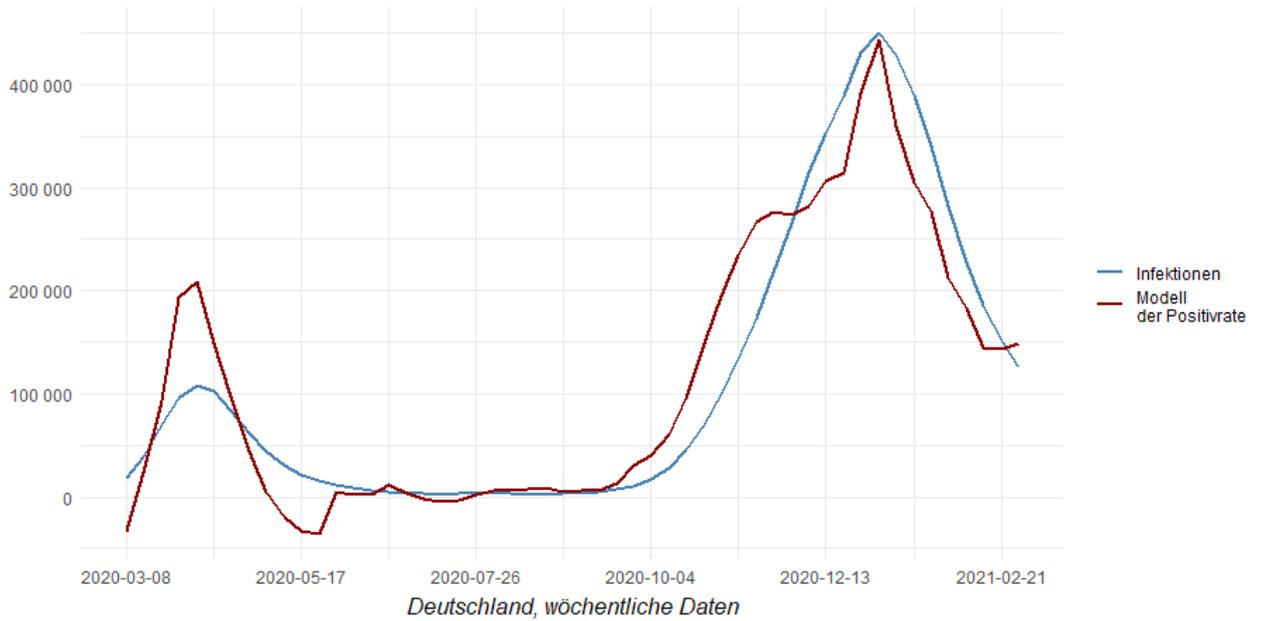


Abbildung 3.25: Modellschätzung durch Positivrate inkl. Strukturbrüche in Deutschland

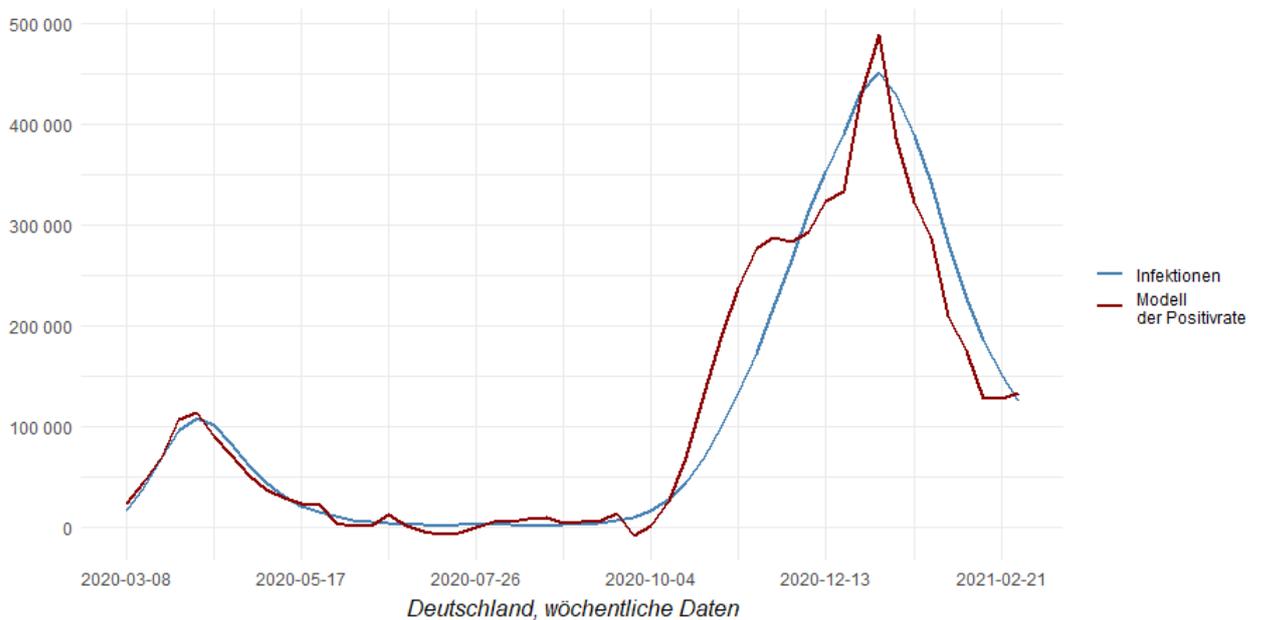


Abbildung 3.26: Modellschätzung durch Positivrate inkl. Strukturbrüche und Interaktionen in Deutschland

## 3.6 Dynamische Modelle

In diesem Kapitel werden dynamische Modelle zur Modellierung des Zusammenhangs zwischen der Positivrate und der Testgröße mit der Dunkelziffer aufgestellt. Wie bereits in Kapitel (2.3) beschrieben, beziehen dynamische Modelle durch die Integration von Lags verzögerte Effekte mit ein. Diese Vorgehensweise wird auf die beiden in den vergangenen Kapiteln als am geeignetsten herausgestellten Modelle, das Testgrößenmodell und das Infektionen-Positivrate-Modell, angewandt. Im ersten Teil dieses Kapitels wird sich mit dem Testgrößenmodell und im zweiten Teil mit dem Positivrate-Infektionen-Modell befasst.

### 3.6.1 Dynamisches Testgrößenmodell

Abbildung (3.21) hat gezeigt, dass beim Testgrößenmodell eine Autokorrelation der Fehler bis zur siebten Ordnung vorzufinden ist. Um dem entgegenzuwirken, wird zunächst ein dynamisches Modell unter Einbeziehung von Lags bis zur fünften Ordnung aufgestellt.

$$\begin{aligned}
 &\text{Anteil der Dunkelziffer}^{(t)} = \\
 &\alpha + \beta_1 \cdot \log(\text{Anteil der Tests})^{(t)} + \beta_2 \cdot \log(\text{Anteil der Tests})^{(t-1)} + \\
 &\beta_3 \cdot \log(\text{Anteil der Tests})^{(t-2)} + \beta_4 \cdot \log(\text{Anteil der Tests})^{(t-3)} + \\
 &\beta_5 \cdot \log(\text{Anteil der Tests})^{(t-4)} + \beta_6 \cdot \log(\text{Anteil der Tests})^{(t-5)} + \epsilon^{(t)} \quad (3.10)
 \end{aligned}$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	0.90950	0.04600	19.774	<2e-16 ***
$\beta_1$	-0.21091	0.10446	-2.019	0.0502
$\beta_2$	0.09146	0.19329	0.473	0.6387
$\beta_3$	0.09564	0.19370	0.494	0.6242
$\beta_4$	-0.11303	0.16545	-0.683	0.4984
$\beta_5$	0.02690	0.14918	0.180	0.8578
$\beta_6$	-0.05029	0.08791	-0.572	0.5705

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.8175$ , adjustiertes  $R^2 = 0.7902$

Tabelle 3.15: Ergebnisse des dynamischen Testgrößenmodells

Die Tabelle (3.15) zeigt die Ergebnisse dieser Modellschätzung. Es zeigt sich, dass bei Einführung der Lags keine der einzelnen Einflussvariablen mehr signifikant ist. Es ist also nicht auszuschließen, dass die einzelnen Lags unter Einbeziehung all der anderen Lags keinen weiteren Einfluss auf die Zielgröße ausübt. Sogar die Richtung des Einflusses variiert von Lag zu Lag, was an den verschiedenen Vorzeichen der Koeffizientenschätzungen zu erkennen ist. Diese Ergebnisse können darauf zurückgeführt werden, dass die sechs Einflussvariablen aus dem obigen Modell sehr stark miteinander korrelieren. Der logarithmierte Anteil der Testgröße zu einem Zeitpunkt  $t$  hängt stark von der Größe dieser Testgröße eine Woche vorher ab, also zum Zeitpunkt  $t - 1$ . Es kommt zu einer hohen Kollinearität zwischen den Einflussvariablen, was zu den unsinnigen Koeffizientenschätzungen aus Tabelle (3.15) führt. Es ist davon auszugehen, dass unter Einbeziehung von fünf Lags, also fünf Wochen, bei einer Gesamtzahl von zweiundfünfzig Wochen, das Modell sich zu stark an unseren Datensatz angepasst hat (*overfitting*), dafür spricht auch das hohe  $R^2$ . Abbildung (3.27) zeigt, dass auch unser Ziel des dynamischen Modells, die Autokorrelation der Fehler zu beheben, verfehlt wurde. Diese besteht immer noch bis zum zweiten Grad. Diese kann wahrscheinlich auch nie eliminiert werden, da der Verlauf der Testgröße über die Zeit hinweg mehrmonatigen Trends unterliegt. In jedem Fall stellt sich heraus, dass dieses dynamische Testgrößenmodell mehr Probleme verursacht als das es löst, somit wird es verworfen.

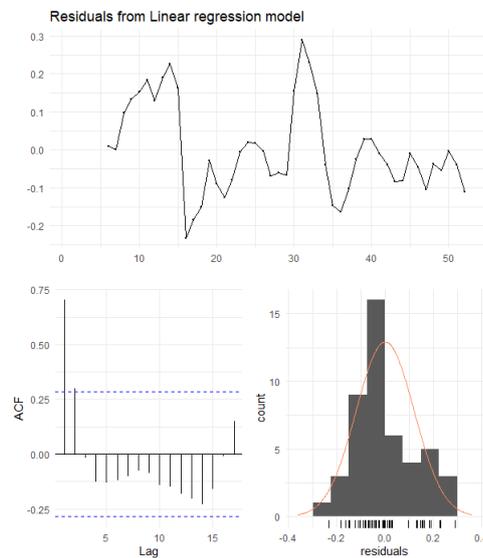


Abbildung 3.27: Residuenanalyse des dynamischen Testgrößenmodells  
*jeweils oben:* Plot der Residuen gegen die Zeit, *jeweils unten links:* ACF-Plot, *jeweils unten rechts:* Häufigkeitsverteilung der Residuen

### 3.6.2 Dynamisches Positivraten-Infektionen-Modell

Abbildung (3.21) hat gezeigt, dass beim Positivraten-Infektionen-Modell eine Autokorrelation der Fehler bis zur dritten Ordnung vorzufinden ist. Um dem entgegenzuwirken, wird analog zum dynamischen Testgrößenmodell zunächst ein dynamisches Modell unter Einbeziehung von Lags bis zur dritten Ordnung aufgestellt.

$$\begin{aligned} \text{Infektionen}^{(t)} = & \\ & \alpha + \beta_1 \cdot \text{Positivrate}^{(t)} + \beta_2 \cdot \text{Positivrate}^{(t-1)} + \\ & \beta_3 \cdot \text{Positivrate}^{(t-2)} + \beta_4 \cdot \text{Positivrate}^{(t-3)} + \epsilon^{(t)} \end{aligned} \quad (3.11)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	-42689	9626	-4.435	6.07e-05 ***
$\beta_1$	26446	7220	3.663	0.000667 ***
$\beta_2$	-3820	13786	-0.277	0.783000
$\beta_3$	-3616	13655	-0.265	0.792358
$\beta_4$	16545	6939	2.384	0.021484 *

signifikant auf einem Niveau von: \*\*\* = 0.001    \*\* = 0.01    \* = 0.05  
 $R^2 = 0.9159$ , adjustiertes  $R^2 = 0.9082$

Tabelle 3.16: Ergebnisse des dynamischen Positivraten-Infektionen-Modells

Das Ergebnis des dynamischen Positivraten-Infektionen-Modells besagt, dass vor allem die kontemporäre Wirkung der Positivrate auf die Infektionen von Bedeutung ist. Jedoch zeigt auch dieses Modell die gleichen Schwächen wie das dynamische Testgrößenmodell. Inhaltlich ergibt es wenig Sinn, warum die Positivrate der letzten und vorletzten Woche keinen Einfluss oder sogar einen negativen Einfluss auf die Infektionszahlen besitzt, der Einfluss der Positivrate von vor drei Wochen hingegen doch signifikant sein soll. Auch hier hat die Kollinearität der Einflussvariablen zu unlogischen Koeffizientenschätzungen geführt. Auch der ACF-Plot aus der Residuenanalyse in Abbildung (3.28) zeigt, dass durch die Integration der Lags die Autokorrelation der Fehler verringert, jedoch nicht aufgehoben werden konnte. Auch eine Erweiterung des Modells auf insgesamt fünf Lags kann zwar die Autokorrelation der Fehler weiter verringern, jedoch nicht beheben. Folglich werden beide dynamischen Modelle verworfen und beide statischen Modelle sind gegenüber den beiden dynamischen Modellen zu

bevorzugen.

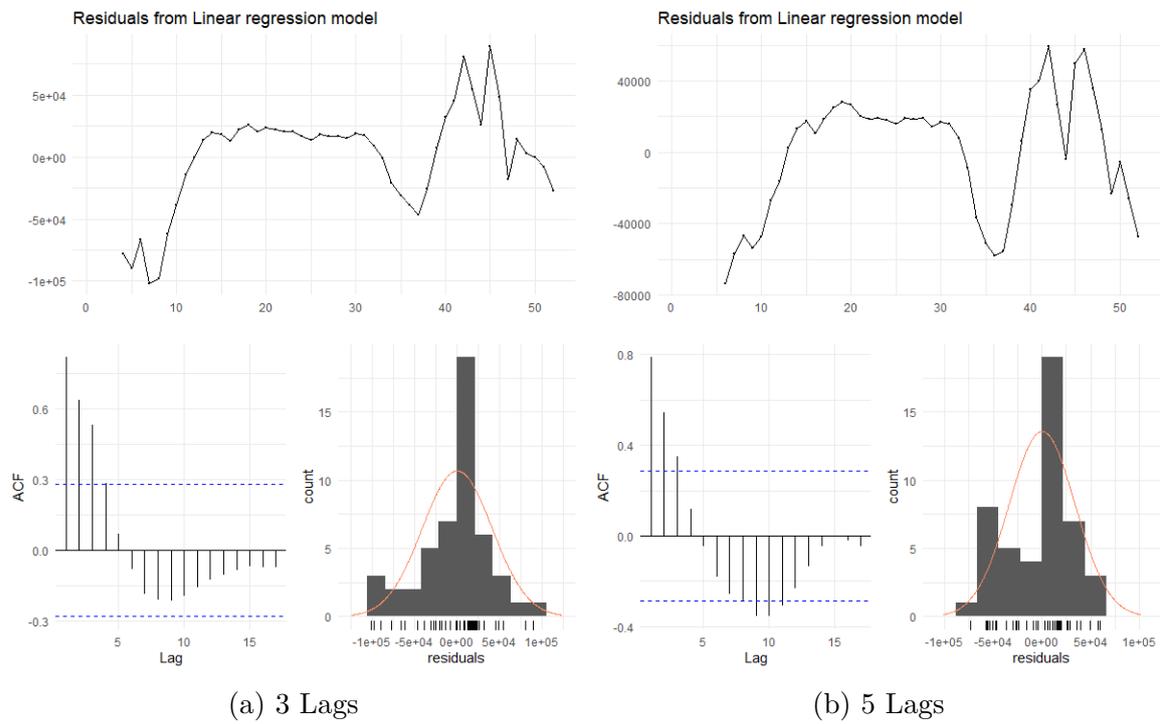


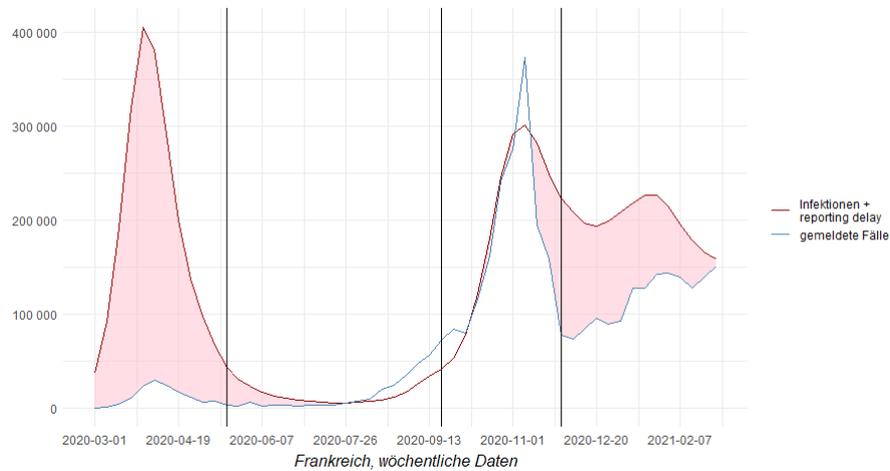
Abbildung 3.28: Residuenanalyse des dynamischen Positivrate-Infektionen-Modells  
*jeweils oben:* Plot der Residuen gegen die Zeit, *jeweils unten links:* ACF-Plot, *jeweils unten rechts:*  
 Häufigkeitsverteilung der Residuen

### 3.7 Betrachtung weiterer europäischer Länder

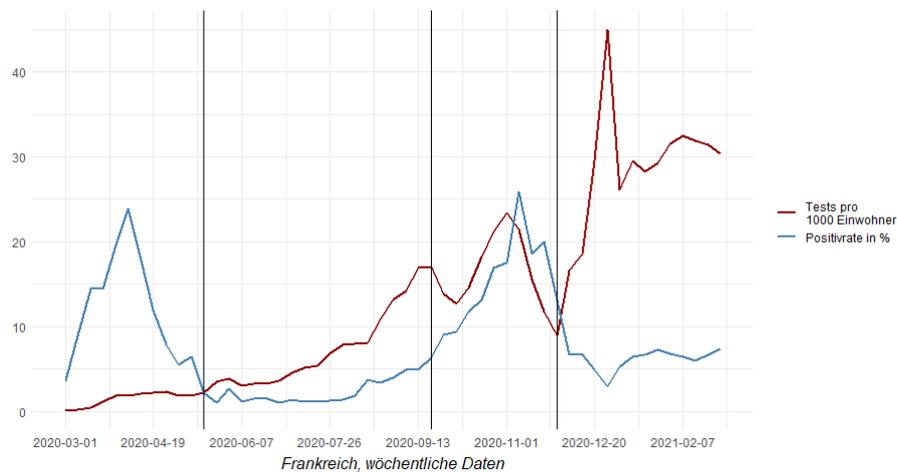
Die gleichen Analysen wie mit Deutschland können auch mit anderen Ländern vorgenommen werden. Im Folgenden werden das statische Positivrate-Infektionen-Modell unter Berücksichtigung von Strukturbrüchen und das statische Testgrößenmodell unter Einbindung der Testgröße und der Dunkelziffer als Anteil an den Infektionen auf die vier Länder Frankreich, Spanien, Italien und Dänemark angewandt. Es werden nur kurz und knapp die wichtigsten Ergebnisse präsentiert. Für das Positivrate-Infektionen-Modell wird zudem nun die Zielvariable *Infektionen* als Infektionen pro 100000 Einwohner betrachtet, um die Koeffizienten besser miteinander vergleichen zu können (für Deutschland z.B. bedeutet diese Transformation eine Multiplikation der Koeffizienten mit dem Faktor  $\frac{100.000}{83.200.000}$ ). Da es sich hier nur um eine lineare Transformation der Zielvariable handelt, beeinflusst das nicht die Modelle.  $R^2$ , Konfidenzintervalle, t-Tests bleiben alle gleich, nur die Koeffizienten verändern sich um einen entsprechenden Faktor. Für die Schätzung der Standardabweichungen der Koeffizientenschätzungen wird die Newey-West-Methode angewandt, da auch bei diesen Ländern von einer Autokorrelation der Fehlerterme auszugehen ist. Dies führt in den meisten Fällen zu einer leichten Verringerung der Signifikanz der Koeffizienten. Für die Schätzung von  $L$  wird dabei die Formel  $L = 0.75 \cdot n^{\frac{1}{3}}$  genutzt, wobei  $n$  für die Anzahl der betrachteten Wochen steht. Um jeweils zu bestimmen an welchen Zeitpunkten Strukturbrüche vorzufinden sind, wird jeweils als allererstes der zeitliche Verlauf der Pandemie mit Fokus auf Infektionszahlen und Testgröße betrachtet.

### 3.7.1 Frankreich

#### Positivrate-Infektionen-Modell



(a) Frankreich: Infektionen und gemeldete Fälle



(b) Frankreich: Testgröße und Positivrate

Abbildung 3.29: Frankreich: Verlauf der Pandemie

Für das Positivrate-Infektionen-Modell wird Frankreich in vier zeitliche Abschnitte unterteilt. Gründe dafür können der Abbildung (3.32) entnommen werden. Frankreich durchläuft im betrachteten Zeitraum zwei große Infektionswellen, einmal zu Beginn der Pandemie und ein zweites Mal ab September 2020. Somit werden hier zwei Strukturbrüche identifiziert, einmal in der Woche vom 17. Mai 2020 und einmal in der Woche vom 20. September 2020 (ersten beide vertikalen Striche). Ein dritter Strukturbruch kann im Verlauf der zweiten Welle gefunden werden, als plötzlich die Testgröße steil

ansteigt (dritter Strich). Einfachheitshalber Wird dieser letzte Zeitabschnitt als dritte Welle bezeichnet, auch wenn die Trennung hier wegen Testgröße gemacht wird. Die *erste Welle* bezeichnet den Abschnitt bis zum ersten Strich, die *zweite Welle* den zwischen dem zweiten und dritten Strich und die *dritte Welle* den Abschnitt nach dem dritten Strich.

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \beta_2 \cdot \text{erste Welle} + \beta_3 \cdot \text{zweite Welle} + \beta_4 \cdot \text{dritte Welle} + \epsilon \quad (3.12)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	-31.50	12.1803	-2.5864	0.01279 *
$\beta_1$	23.81	2.1779	10.9315	1.266e-14 ***
$\beta_2$	40.83	26.3347	1.5505	0.12759
$\beta_3$	-48.33	34.7861	-1.3892	0.17117
$\beta_4$	172.39	24.0209	7.1767	3.943e-09 ***

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.9136$ , adjustiertes  $R^2 = 0.9064$

Tabelle 3.17: Frankreich: Ergebnisse des Positivrate-Infektionen-Modells

Die Ergebnisse aus Tabelle (3.17) zeigen, dass auch hier die zeitliche Trennung signifikant ist, sowie der Einfluss der Positivrate. Im Mittel bringt zu jedem Zeitpunkt eine Erhöhung der Positivrate um einen Prozent eine Erhöhung der Infektionen pro 100000 Einwohner um rund 24 Infektionen nach sich. Dabei ist z.B. zur Zeit hoher Infektionen und sehr hoher Testzahlen (dritte Welle), bei gleichbleibender Positivrate, im Mittel davon auszugehen, dass die Zahl der Infektionen um rund  $172 + 48 = 220$  Infektionen pro 100000 Menschen größer ist als zur Zeit der zweiten Wellen, wo die Zahl der Infektionen ungefähr gleich groß ist, aber die Zahl der Tests deutlich kleiner.

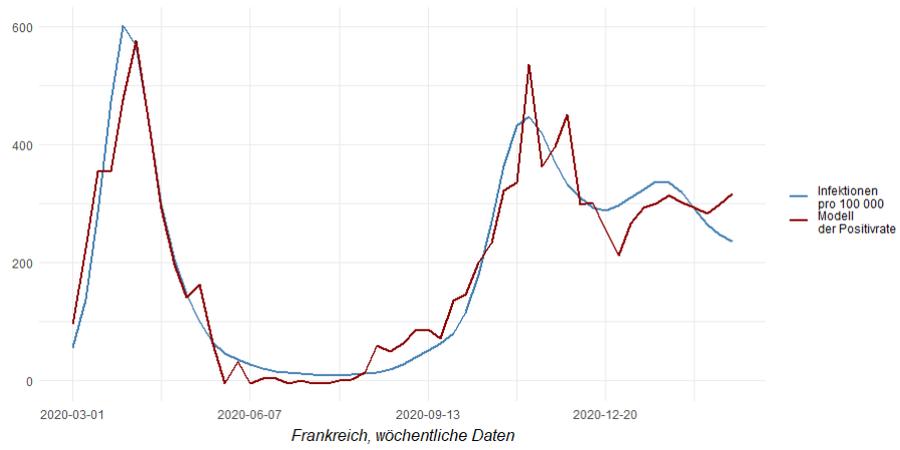


Abbildung 3.30: Modellschätzung der Infektionen durch Positivrate in Frankreich

### Testgrößenmodell

$$\text{Anteil der Dunkelziffer} = \alpha + \beta_1 \cdot \log(\text{Anteil der Tests}) + \epsilon \quad (3.13)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	0.74025	0.062732	11.8002	3.399e-16 ***
$\beta_1$	-0.15119	0.021128	-7.1558	3.089e-09 ***

signifikant auf einem Niveau von: \*\*\* = 0.001    \*\* = 0.01    \* = 0.05  
 $R^2 = 0.5144$ , adjustiertes  $R^2 = 0.5049$

Tabelle 3.18: Frankreich: Ergebnisse des Testgrößenmodells

Der Einfluss der Testgröße ist signifikant. Erhöht sich der relative Anteil der Tests um den Faktor 2,72, so verringert sich die relative Dunkelziffer im Mittel um 0,15.

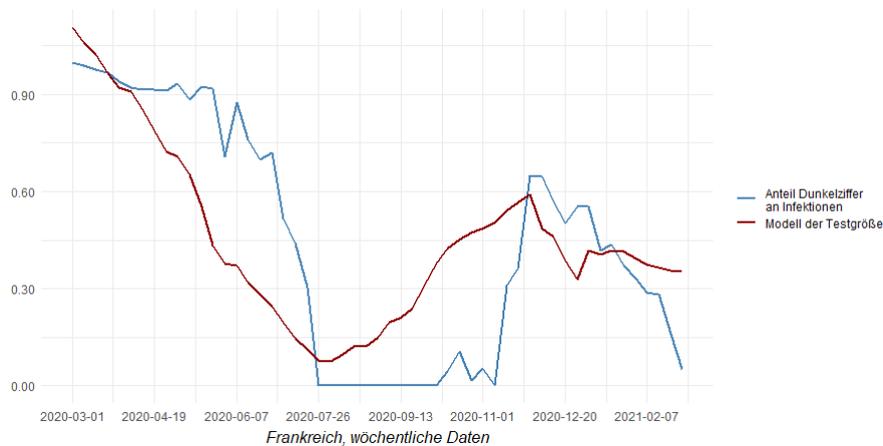
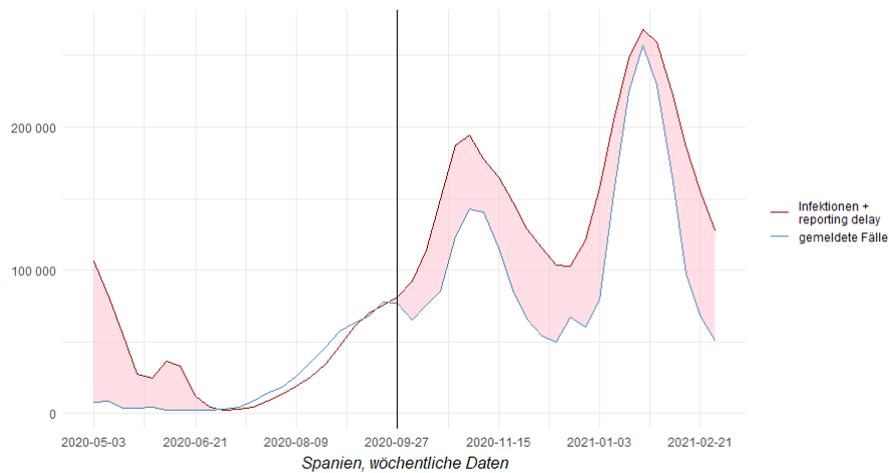


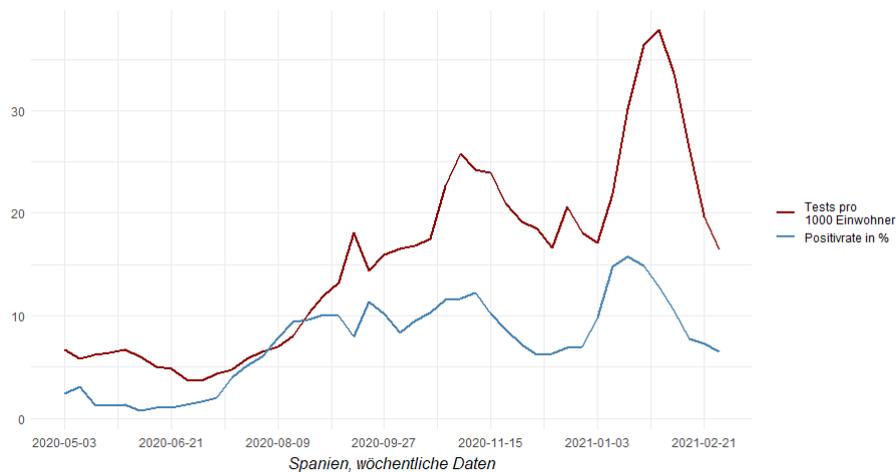
Abbildung 3.31: Modellschätzung durch Testgröße in Frankreich

## 3.7.2 Spanien

### Positivrate-Infektionen-Modell



(a) Spanien: Infektionen und gemeldete Fälle



(b) Spanien: Testgröße und Positivrate

Abbildung 3.32: Spanien: Verlauf der Pandemie

Für Spanien wird nur eine zeitliche Trennung vorgenommen, da es weniger radikale Einschnitte der Variablen gibt, sondern eher graduelle Veränderungen. Die Trennung wird vorgenommen um eine zeitliche Trennung zwischen niedrigen Infektionszahlen zu Beginn der Pandemie und stark steigenden nach der Trennung zu kreieren. Die Zeit nach der Trennung wird als *erste Welle* bezeichnet.

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \beta_2 \cdot \text{erste Welle} + \beta_3 \cdot \text{erste Welle} \cdot \text{Positivrate} + \epsilon \quad (3.14)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	51.124	32.8597	1.5558	0.1276
$\beta_1$	5.675	4.2038	1.3500	0.1846
$\beta_2$	-12.302	50.7017	-0.2426	0.8095
$\beta_3$	25.897	5.2755	4.9089	1.582e-05 ***

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.8669$ , adjustiertes  $R^2 = 0.8569$

Tabelle 3.19: Spanien: Ergebnisse des Positivrate-Infektionen-Modells

Hier ist die Positivrate nur zur Zeit der ersten Welle signifikant. Eine Erhöhung der Positivrate um einen Prozent hat während der ersten Welle eine mittlere Erhöhung der Infektionen pro 100000 Einwohner um rund  $6 + 26 = 32$  zur Folge. Vor Beginn der ersten Welle, also zur Zeit weniger Infektionen und weniger Tests, lässt sich die Zahl der Infektionen nicht gut mithilfe der Positivrate schätzen, was auch in Abbildung (3.33) zu erkennen ist.

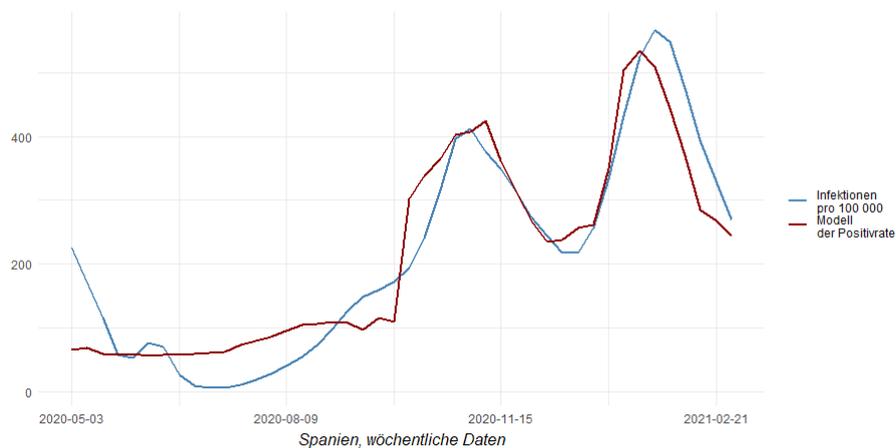


Abbildung 3.33: Modellschätzung der Infektionen durch Positivrate in Spanien

### Testgrößenmodell

$$\text{Anteil der Dunkelziffer} = \alpha + \beta_1 \cdot \log(\text{Anteil der Tests}) + \epsilon \quad (3.15)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	0.81510	0.168818	4.8282	1.856e-05 ***
$\beta_1$	-0.20313	0.063661	-3.1909	0.002684 **

signifikant auf einem Niveau von: \*\*\* = 0.001    \*\* = 0.01    \* = 0.05  
 $R^2 = 0.2176$ , adjustiertes  $R^2 = 0.199$

Tabelle 3.20: Spanien: Ergebnisse des Testgrößenmodells

Der Einfluss der Testgröße ist signifikant. Erhöht sich der relative Anteil der Tests um den Faktor 2,72, so verringert sich die relative Dunkelziffer im Mittel um 0,2.

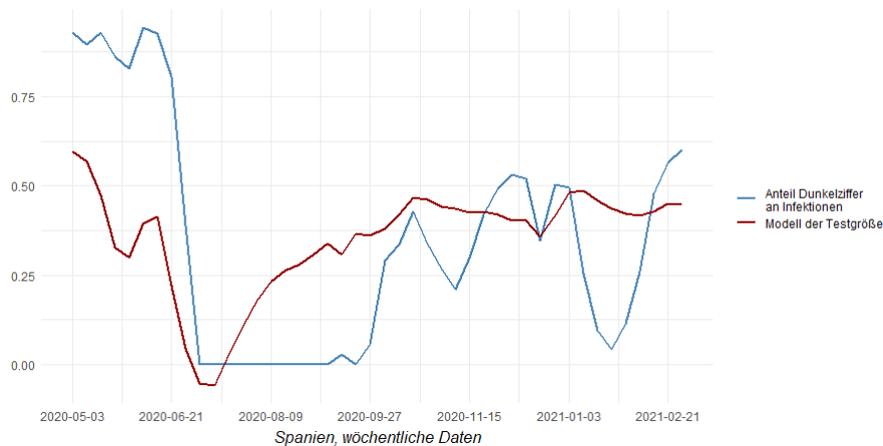
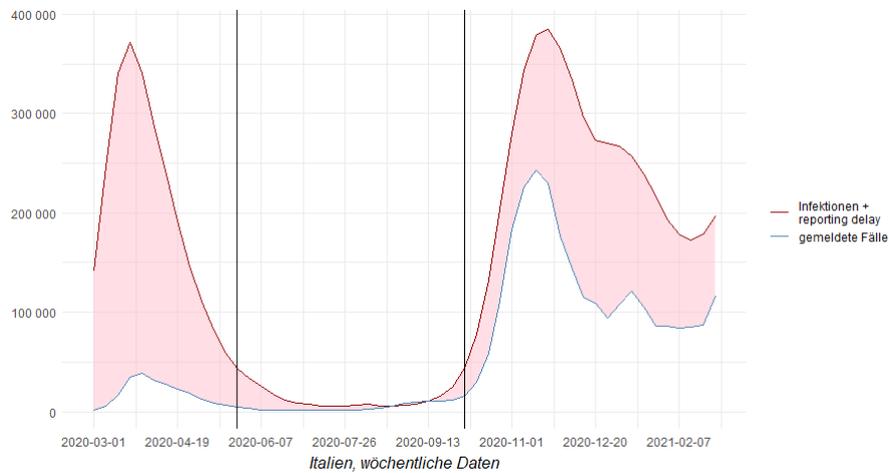


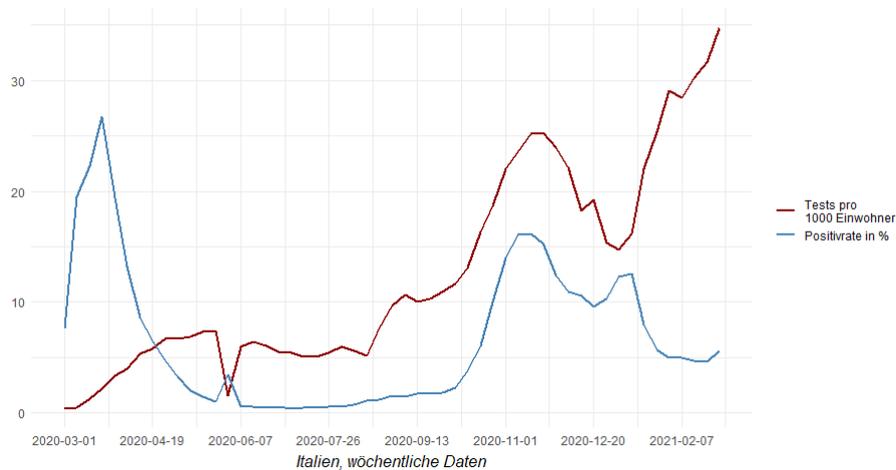
Abbildung 3.34: Modellschätzung durch Testgröße in Spanien

### 3.7.3 Italien

#### Positivrate-Infektionen-Modell



(a) Italien: Infektionen und gemeldete Fälle



(b) Italien: Testgröße und Positivrate

Abbildung 3.35: Italien: Verlauf der Pandemie

In Italien werden ähnlich wie in Deutschland zwei zeitliche Trennungen vorgenommen, um die Perioden der *ersten Welle*, der ruhigen Sommermonate dazwischen und der *dritten Welle* zu trennen. Während der ersten und der zweiten Welle sind die Infektionszahlen ungefähr gleich groß, allerdings werden zur Zeit der zweiten Welle viel mehr Tests durchgeführt.

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \beta_2 \cdot \text{erste Welle} + \beta_3 \cdot \text{zweite Welle} + \epsilon \quad (3.16)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	-5.333	6.1943	-0.8609	0.39347
$\beta_1$	23.657	2.9554	8.0049	1.878e-10 ***
$\beta_2$	94.012	35.1174	2.6771	0.01007 *
$\beta_3$	192.436	39.0464	4.9284	9.886e-06 ***

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.922$ , adjustiertes  $R^2 = 0.9173$

Tabelle 3.21: Italien: Ergebnisse des Positivrate-Infektionen-Modells

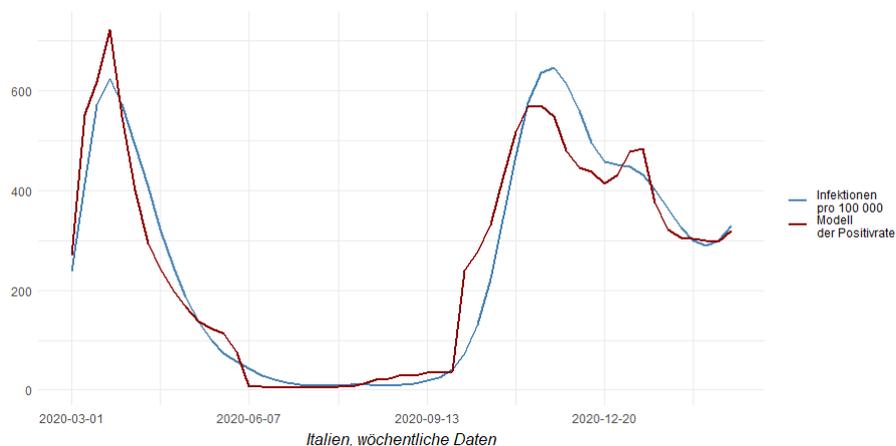


Abbildung 3.36: Modellschätzung der Infektionen durch Positivrate in Italien

Hier sind alle Koeffizienten signifikant. Eine Erhöhung der Positivrate um einen Prozent führt zu einer mittleren Erhöhung der Infektionen pro 100000 Einwohner um rund 24. Interessant zu sehen ist, dass bei gleichbleibender Positivrate die Zahl der Infektionen während der zweiten Welle, wenn die Infektionszahlen ungefähr gleich groß sind wie während der ersten Welle, jedoch die Anzahl der Tests deutlich größer, die Zahl der Infektionen, bei gleichbleibender Positivrate, im Mittel um rund  $192 - 94 = 98$  höher zu schätzen ist.

**Testgrößenmodell**

$$\text{Anteil der Dunkelziffer} = \alpha + \beta_1 \cdot \log(\text{Anteil der Tests}) + \epsilon \quad (3.17)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	0.79588	0.056100	14.1869	< 2e-16 ***
$\beta_1$	-0.08290	0.033304	-2.4892	0.01611 *

signifikant auf einem Niveau von: \*\*\* = 0.001    \*\* = 0.01    \* = 0.05  
 $R^2 = 0.2686$ , adjustiertes  $R^2 = 0.2543$

Tabelle 3.22: Italien: Ergebnisse des Testgrößenmodells

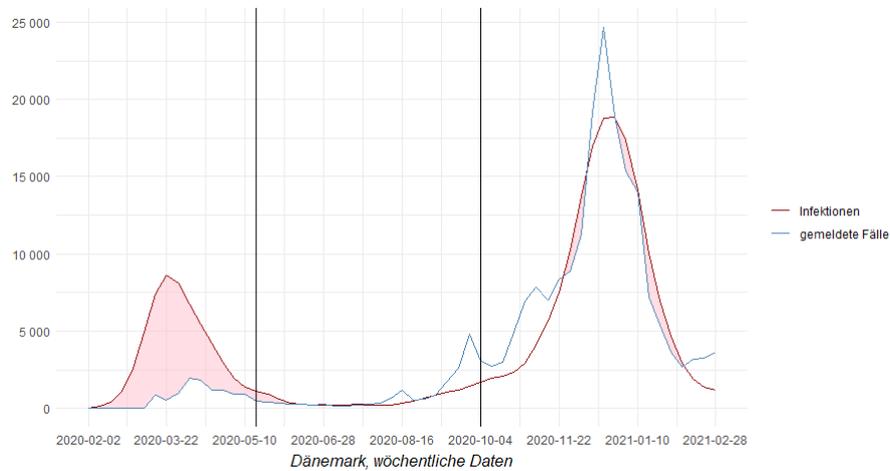
Der Einfluss der Testgröße ist signifikant. Erhöht sich der relative Anteil der Tests um den Faktor 2,72, so verringert sich die relative Dunkelziffer im Mittel um 0,08.



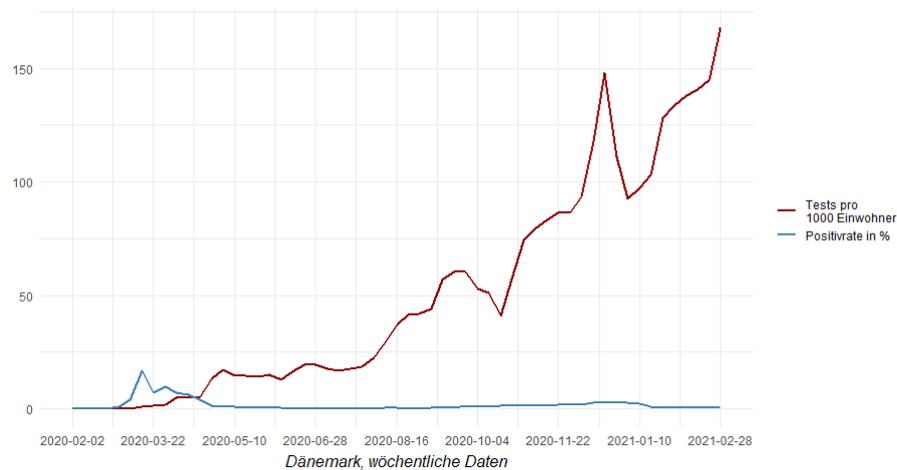
Abbildung 3.37: Modellschätzung durch Testgröße in Italien

### 3.7.4 Dänemark

#### Positivrate-Infektionen-Modell



(a) Dänemark: Infektionen und gemeldete Fälle



(b) Dänemark: Testgröße und Positivrate

Abbildung 3.38: Dänemark: Verlauf der Pandemie

Auch in Dänemark wird die Pandemie in drei Abschnitte unterteilt, einer ersten Welle hoher Infektionszahlen vor dem ersten vertikalen Strich und einer zweiten solchen Welle nach dem zweiten vertikalen Strich.

$$\text{Infektionen} = \alpha + \beta_1 \cdot \text{Positivrate} + \beta_2 \cdot \text{erste Welle} + \beta_3 \cdot \text{zweite Welle} + \beta_4 \cdot \text{zweite Welle} \cdot \text{Positivrate} + \epsilon \quad (3.18)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	5.367	2.0013	2.6819	0.009789 **
$\beta_1$	9.353	1.9861	4.7090	1.889e-05 ***
$\beta_2$	20.278	10.0999	2.0078	0.049874 *
$\beta_3$	-34.801	33.3710	-1.0428	0.301848
$\beta_4$	98.504	15.4585	6.3722	4.951e-08 ***

signifikant auf einem Niveau von: \*\*\* = 0.001 \*\* = 0.01 \* = 0.05  
 $R^2 = 0.8491$ , adjustiertes  $R^2 = 0.8375$

Tabelle 3.23: Dänemark: Ergebnisse des Positivrate-Infektionen-Modells

Die Positivrate und die zeitliche Trennung sind signifikant. Eine Erhöhung der Positivrate um einen Prozent hat im Mittel während der ersten Welle eine Erhöhung der Infektionszahlen um rund 9 Infektionen pro 100000 Einwohner zur Folge. Während der zweiten Welle, wenn deutlich mehr Tests durchgeführt werden, sorgt eine Erhöhung der Positivrate um einen Prozent für eine mittlere Erhöhung der Infektionen pro 100000 Einwohner um ganze  $9 + 99 = 108$  Infektionen.

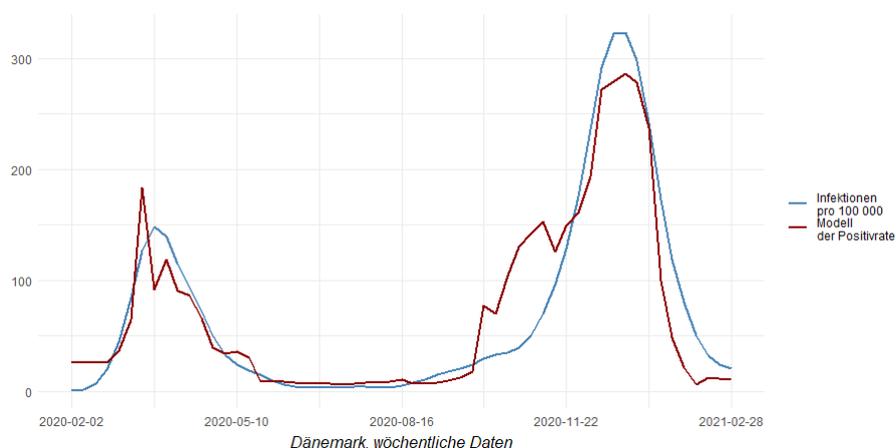


Abbildung 3.39: Modellschätzung der Infektionen durch Positivrate in Dänemark

### Testgrößenmodell

$$\text{Anteil der Dunkelziffer} = \alpha + \beta_1 \cdot \log(\text{Anteil der Tests}) + \epsilon \quad (3.19)$$

Koeffizient	Schätzung	Standardabweichung	t-Wert	P(> t )
$\alpha$	0.80159	0.0569102	14.085	< 2.2e-16 ***
$\beta_1$	-0.12029	0.0099169	-12.130	< 2.2e-16 ***

signifikant auf einem Niveau von: \*\*\* = 0.001    \*\* = 0.01    \* = 0.05  
 $R^2 = 0.6996$ , adjustiertes  $R^2 = 0.6942$

Tabelle 3.24: Dänemark: Ergebnisse des Testgrößenmodells

Der Einfluss der Testgröße ist signifikant. Erhöht sich der relative Anteil der Tests um den Faktor 2,72, so verringert sich die relative Dunkelziffer im Mittel um 0,12.

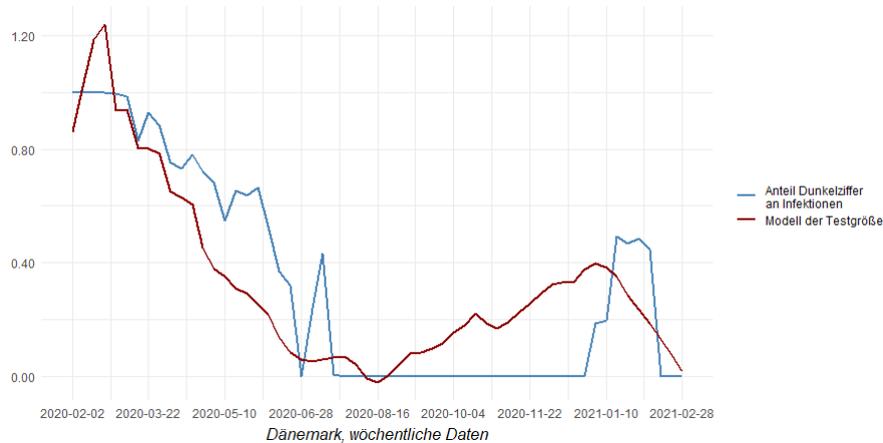


Abbildung 3.40: Modellschätzung durch Testgröße in Dänemark

# 4 Diskussion

## 4.1 Zusammenfassende Bewertung

### 4.1.1 Zusammenfassung der Modellergebnisse

Zusammenfassend kann festgehalten werden, dass sowohl für Deutschland, als auch für die Länder Frankreich, Spanien, Italien und Dänemark, sehr gute Modelle für die Modellierung des Zusammenhangs zwischen Testpositivrate und Infektionen und gute Modelle für die Modellierung des Zusammenhanges zwischen relativer Testgröße und relativer Dunkelziffer für den Zeitraum zwischen Februar 2020 und Februar 2021 aufgestellt werden können. Dabei besitzt die Positivrate stets in irgendeiner Form einen signifikanten, positiven Einfluss auf die Zahl der Infektionen eines Landes. Die relative Testgröße besitzt hingegen stets einen negativen signifikanten Einfluss auf den Anteil nicht entdeckter Fälle an allen Infektionen. Eine Erhöhung der relativen Testgröße um den Faktor 2,72 führt für unsere betrachteten Länder zu einer mittleren Reduktion des Anteils unentdeckter Fälle an allen Infektionen zwischen 8% (Italien) und 20% (Spanien). Eine Änderung der Positivrate um einen Prozent, führt in Frankreich und Italien zu einer mittleren Erhöhung der Infektionszahlen von rund 24 Infektionen pro 100000 Einwohnern und in Spanien zu einer mittleren Erhöhung von rund 32 Infektionen pro 100000 Einwohnern. In Dänemark und Deutschland verhält sich hingegen der Einfluss der Positivrate unterschiedlich zu verschiedenen Perioden der Pandemie. Während der ersten Infektionswelle führt in Deutschland eine Erhöhung der Positivrate um einen Prozent zu einer mittleren Erhöhung der Infektionszahlen um rund 14 Infektionen pro 100000 Einwohnern, in Dänemark um rund 9 Infektionen pro 100000 Einwohnern. Hingegen führt, während der zweiten Infektionswelle, eine Erhöhung der Positivrate um einen Prozent im Mittel zu einer Erhöhung der Infektionszahlen um rund 44 (Deutschland) und ganzen 108 (Dänemark) Infektionen pro 100000 Einwohnern. Bei der Modellierung mithilfe der Positivrate ist zu beachten, dass die Einführung von verschiedenen *intercepts* für unterschiedliche Zeiträume entschieden zur Verbesserung der Modelle beiträgt. Die zeitlichen Abschnitte sollten dabei abhängig von der Dynamik des Infektionsgeschehens und des Ausmaßes an durchgeführten Tests gewählt werden.

### 4.1.2 Zusammenfassung der methodischen Hintergründe

Diese Arbeit zeigt, dass sich dynamische Modelle für die Analyse des Zusammenhangs von Dunkelziffer, Positivrate und Testgröße als ungeeignet herausstellen. Die statischen Modelle weisen jedoch auch Probleme auf, da sie die zentralen Anforderungen bei der Regressionsanalyse von Zeitreihen, nämlich die Stationarität der Variablen und das Ausschließen von Autokorrelation der Fehler, verletzen. Mit der Autokorrelation der Fehler kann durch die Neuschätzung der Standardabweichungen der Koeffizienten mithilfe der Newey-West-Methode gut umgegangen werden. Das Problem der Stationarität der Variablen kann durch die Neuspezifikation der Modellgleichungen mithilfe von Differenzbildungen nicht behoben werden, da diese Methode durch die Eliminierung langfristiger Trends der Variablen auch langfristige Zusammenhänge eliminiert. Letztendlich kann, durch die Berücksichtigung von Strukturbrüchen in den Modellgleichungen, die Stationarität der Variablen berücksichtigt werden.

Außerdem zeigt diese Arbeit, dass eine naive Herangehensweise zur Modellierung des Zusammenhangs zwischen Dunkelziffer, Testgröße und Positivrate nicht geeignet ist. Dies liegt an der zeitlichen Korrelation der Variablen. Infektionszahlen, Testgröße und gemeldete Fallzahlen weisen im Verlauf der Pandemie gemeinsame Trends auf, die eher zufällig sind oder vielmehr auf politische Maßnahmen zur Bekämpfung der Pandemie fußen. Diesen Zusammenhängen der Variablen muss in der Aufstellung von Modellgleichungen Rechnung getragen werden.

## 4.2 Ausblick

Aufbauend auf den Ergebnissen dieser Arbeit, können weiterführende Überlegungen zum Zusammenhang zwischen Positivrate, Infektionen, Dunkelziffer und Testgröße aufgestellt werden. Es wäre interessant zu analysieren, wie genau die Dynamik der Infektionszahlen und der Testgröße die Testpositivrate beeinflusst. In dieser Arbeit wurde die Trennung der Modelle in mehrere Zeitabschnitte nur mit dem Augenschein vorgenommen. Womöglich kann man diese Trennung auch mathematisch begründen, dass bspw. bei einem bestimmten Verhältnis von Infektionszahlen und Testgröße, der Einfluss der Positivrate in einem bestimmten Intervall liegt. Dann könnte man diese Ergebnisse auch für die Vorhersage von Infektionszahlen nutzen. Kennt man dann die Infektionszahlen, ließe sich dann auch, mit den in dieser Arbeit aufgestellten Testgrößenmodelle, der Einfluss einer Erhöhung durchgeführter Tests auf die Dunkelziffer schätzen. Bezieht man noch mehr Länder in die Analyse mit ein, könnte man womöglich Länder und Zeiträume mit einer ähnlichen Dynamik von Infektionszahlen und Tests im Verhältnis zur Bevölkerungszahl finden. Dann könnte man sehen, ob der Einfluss der Änderung der Positivrate auf die Infektionen auch gleich aussieht. So könnte womöglich ein Modell aufgestellt werden, das den Einfluss der Positivrate auf die Infektionszahlen unter Einbeziehung der Testgröße beschreibt.

An dieser Stelle sollte erwähnt werden, dass das Positivrate-Infektionen-Modell auch ohne die Berücksichtigung der Strukturbrüche für alle Länder, außer Dänemark, ziemlich gut funktioniert. Somit können auch damit Prognosen für die Infektionszahlen nur mithilfe der Positivrate aufgestellt werden. Man könnte sich auch anschauen, wenn bspw. in einem Land die Infektionszahlen wieder steigen, wie in diesem Land zu einem früheren Zeitpunkt, bspw. einer letzten Infektionswelle, die Positivrate die Zahl der Infektionen beeinflusst hat und dieses Modell auf die Gegenwart beziehen. Oder man könnte sich vergleichbare Länder anschauen, die zu einem früheren Zeitpunkt eine ähnliche Situation hatten, und sich da den Einfluss der Positivrate auf die Infektionen anschauen.

Ein weiterer Aspekt, der bei der Analyse des Zusammenhangs von Dunkelziffer, Testgröße und Positivrate berücksichtigt werden könnte, sind politische Entscheidungen, die den Umgang der Menschen mit Tests beeinflussen. Wie könnte die Pflicht zum Testen nach Urlaubsreisen die Zahl der durchgeführten Tests und die Positivrate beeinflussen?

Vielleicht lassen sich dann mehr gesunde Menschen testen und die Positivrate sinkt, obwohl die Zahl der Infektionen gleich bleibt. Oder wie hat sich das massenhafte Angebot von kostenlosen Schnelltests auf die Testgröße und die Positivrate ausgewirkt? Vielleicht lassen sich dann nur die Menschen testen, die zuerst einen positiven Schnelltest hatten. Dies könnte z.B. die Positivrate eher positiv beeinflussen.

### 4.2.1 Spezifität und Sensitivität

Einen interessanten Aspekt bei der Definition der Dunkelziffer bildet die Rolle der Sensitivität und Spezifität von Tests. Abbildung (4.1) zeigt den Einfluss von unterschiedlichen Werten für eine Sensitivität und eine Spezifität ungleich eins, während den Sommermonaten in Deutschland im Jahre 2020. Man sieht deutlich einen großen Effekt auf die gemeldeten Fallzahlen, vor allem, wenn die Sensitivität als eher hoch und die Spezifität als niedrig angenommen wird. Die Zahl falsch positiver Testergebnisse ist in Deutschland im Sommer 2020 besonders hoch, da hier die Infektionszahlen besonders niedrig sind, jedoch die Zahl durchgeführter Tests trotzdem sehr hoch. Vor allem die rechte Grafik zeigt besonders gut diesen starken Effekt, hier wird von einer Sensitivität von 0.9 ausgegangen. Man sieht, dass, in der Mitte zwischen dem 21. Juni und dem 9. August, alle knapp 2500 gemeldeten Fälle bei einer Spezifität von 0.995 auf falsch positive Ergebnisse zurückgeführt werden können. Auch in den Wochen davor und danach können meist um die Hälfte aller gemeldeten Fallzahlen auf falsch positive Ergebnisse zurückgeführt werden. Dies könnte erklären, warum in dieser Arbeit für die Sommermonate in Deutschland zunächst eine negative Dunkelziffer berechnet wurde, die dann korrigiert werden musste. Die linke Grafik zeigt, dass eine verringerte Sensitivität den Effekt der Spezifität wieder ausgleichen würde.

Abbildung (4.2) zeigt den gegenteiligen Effekt während der Wintermonate zwischen den Jahren 2020 und 2021 in Deutschland. In dieser Periode sehr hoher Infektionszahlen kommt der Effekt von Werten für eine Sensitivität ungleich eins zum Vorschein. Es werden zwischen November und Januar wöchentlich um die 50000 infizierte Menschen fälschlicherweise als gesund angesehen, bei einer Zahl zwischen 175000 und 250000 Infizierten, die sich testen lassen. Im Hinblick auf die Dunkelziffer kann man sagen, dass sie womöglich aufgrund falsch negativer Ergebnisse zu hoch ausfällt.

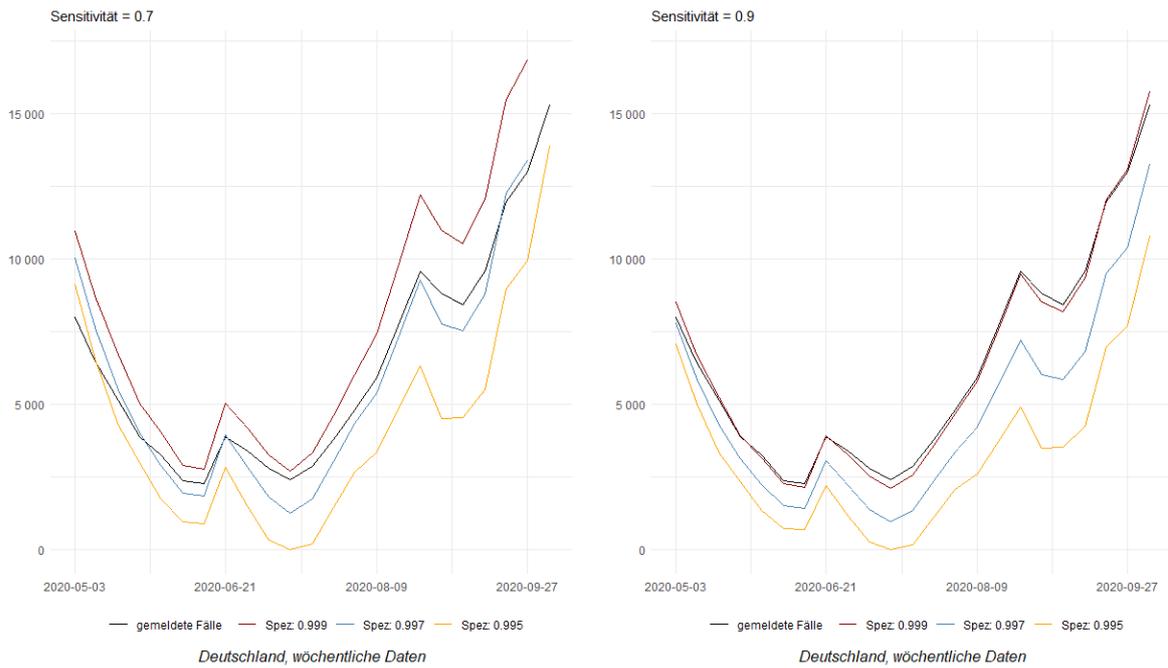


Abbildung 4.1: Einflüsse von Sensitivität und Spezifität in den Sommermonaten in Deutschland

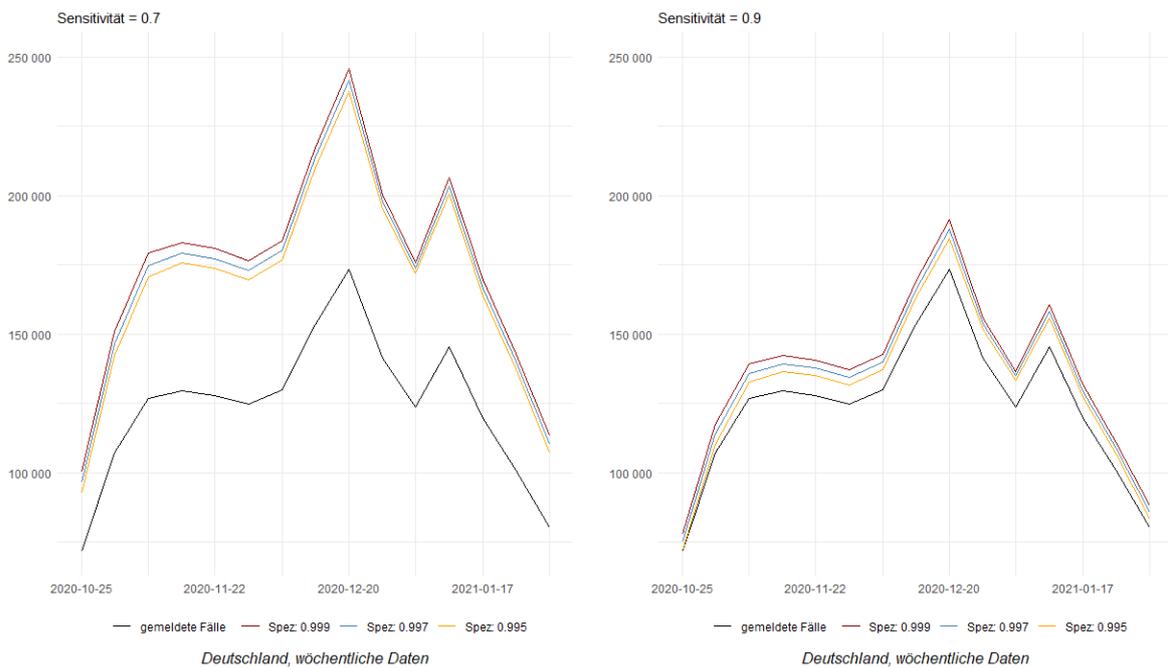


Abbildung 4.2: Einflüsse von Sensitivität und Spezifität während der zweiten Welle in Deutschland

# Literaturverzeichnis

- Baden-Württemberg. *Lockerungen mit vier Inzidenzstufen*. letzter Zugriff am 28. Juni 2021, 2021. URL [https://www.baden-wuerttemberg.de/fileadmin/redaktion/dateien/PDF/Coronainfos/210625\\_Auf\\_einen\\_Blick\\_DE.pdf](https://www.baden-wuerttemberg.de/fileadmin/redaktion/dateien/PDF/Coronainfos/210625_Auf_einen_Blick_DE.pdf).
- ECDC. *About ECDC*. letzter Zugriff am 05. Mai 2021, 2021a. URL <https://www.ecdc.europa.eu/en/about-ecdc>.
- ECDC. *Download historical data (to 14 December 2020) on the daily number of new reported COVID-19 cases and deaths worldwide*. letzter Zugriff am 02. Mai 2021, 2021b. URL <https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide>.
- ECDC. *How ECDC collects and processes COVID-19 data*. letzter Zugriff am 24. Juni 2021, 2021c. URL <https://www.ecdc.europa.eu/en/covid-19/data-collection>.
- ECDC. *Data on testing for COVID-19 by week and country*. letzter Zugriff am 20. Mai 2021, 2021d. URL <https://www.ecdc.europa.eu/en/publications-data/covid-19-testing>.
- ECDC. *Data dictionary on COVID-19 testing*. letzter Zugriff am 20. Juni 2021, 2021e. URL [https://www.ecdc.europa.eu/sites/default/files/documents/2021-01-13\\_Variable\\_Dictionary\\_and\\_Disclaimer\\_weekly\\_testing\\_data\\_EUEEAUK.pdf](https://www.ecdc.europa.eu/sites/default/files/documents/2021-01-13_Variable_Dictionary_and_Disclaimer_weekly_testing_data_EUEEAUK.pdf).
- ECDC. *ECDC's mission and main activities*. letzter Zugriff am 10. Mai 2021, 2021f. URL <https://www.ecdc.europa.eu/en/about-us/ecdc-mission-and-main-activities>.
- S. Flaxman, S. Mishra, A. Gandy, et al. *Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe*. *Nature* 584, S.257–261, 2020. doi: 10.1038/s41586-020-2405-7. URL <https://doi.org/10.1038/s41586-020-2405-7>.
- GitHub. *socialmixr*. letzter Zugriff am 28. April 2021, 2021a. URL <https://github.com/sbfknk/socialmixr>.

- GitHub. *CSSEGISandData/COVID-19*. letzter Zugriff am 26. April 2021, 2021b. URL [https://github.com/CSSEGISandData/COVID-19/blob/master/csse\\_covid\\_19\\_data/csse\\_covid\\_19\\_time\\_series/time\\_series\\_covid19\\_confirmed\\_global.csv](https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv).
- Felix Günther, Ursula Berger, Michael Höhle, Andreas Bender, Manfred Wildner, Iris M. Heid, und Helmut Küchenhoff. *Analysis of COVID-19 case numbers: adjustment for diagnostic misclassification on the example of German case reporting data*. Cold Spring Harbor Laboratory Press, medRxiv, 2021. doi: 10.1101/2020.09.29.20203877. URL <https://www.medrxiv.org/content/early/2021/03/10/2020.09.29.20203877.full.pdf>.
- JHU.edu. *COVID-19 Dashboard*. letzter Zugriff am 27. Juni 2021, 2021. URL <https://coronavirus.jhu.edu/map.html>.
- Joël Mossong, Niel Hens, Mark Ji, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, et al. *Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases*. PLOS MEDICINE, 2008. URL <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0050074#s3>.
- F.W. Schwartz, B.-P. Robra, und J.G. Brecht. *Prävention und Früherkennung des kolorektalen Karzinoms*. Springer, Berlin, Heidelberg, 1984. ISBN 978-3-540-12865-6. doi: 10.1007/978-3-642-69388-5\_5. URL [https://doi.org/10.1007/978-3-642-69388-5\\_5](https://doi.org/10.1007/978-3-642-69388-5_5).
- J. H. Stock und M. W. Watson. *Introduction to econometrics*. Amsterdam: Pearson, 2015.
- Matthias-W. Stoetzer. *Regressionsanalyse in der empirischen Wirtschafts- und Sozialforschung Band 2 Komplexe Verfahren*. Springer, Berlin, Heidelberg, 2020. doi: 978-3-662-61437-2.
- UnitedNations. *World Population Prospects 2019*. Medium variant, annual, from 1950 to 2100 (CSV, 113.05 MB), letzter Zugriff am 24. März 2021, 2019. URL <https://population.un.org/wpp/Download/Standard/CSV/>.
- R. Verity, L. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, et al. *Estimates of the severity of coronavirus disease 2019: a model-based ana-*

*lysis*. Lancet Infect Dis, Vol.20, No.6, S.669-667, 2020. doi: 10.1016/S1473-3099(20)30243-7. URL [https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(20\)30243-7/fulltext#](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30243-7/fulltext#).

P.G.T. Walker, C. Whittaker, O. Watson, et al. *The Global Impact of COVID-19 and Strategies for Mitigation and Suppression*. 2020. doi: 10.25561/77735. URL <https://spiral.imperial.ac.uk/bitstream/10044/1/77735/10/2020-03-26-COVID19-Report-12.pdf>.

Jessica Watson, Penny F Whiting, und John E Brush. *Interpreting a covid-19 test result*. British Medical Journal, 2020. URL <https://www.bmj.com/content/369/bmj.m1808/>.