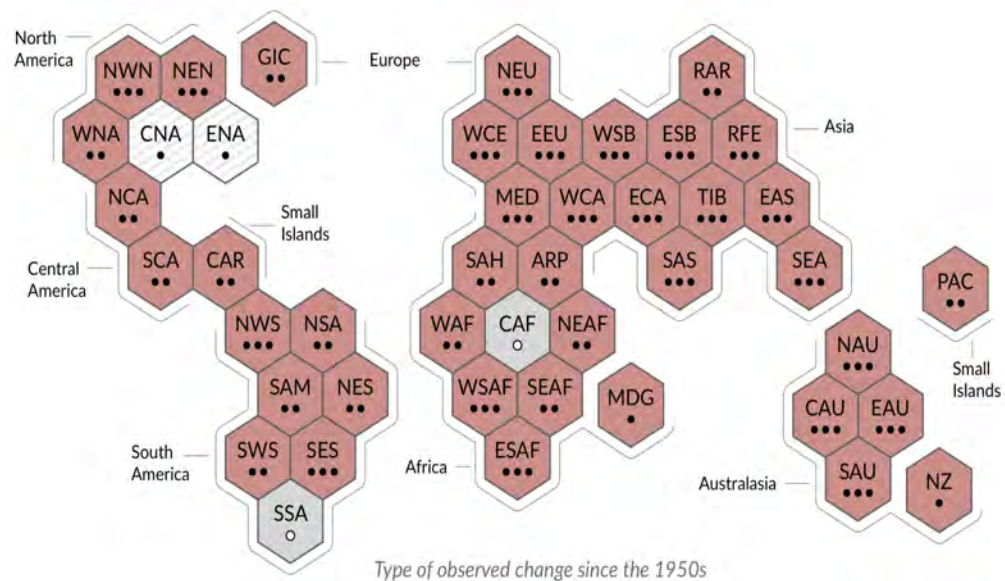


# Regionalökonomie I

– 01.03. 2024 –



Kristof Dascher<sup>1</sup>

<sup>1</sup>Dieses Skript ist für Sie persönlich gedacht; bitte nutzen Sie dieses Skript nur selbst, machen Sie es insbes. bitte nicht öffentlich. Das Titelbild ist aus: IPCC (2020) Anthropomorphe Klimaerwärmung. Im Text finden Sie viele Querverweise auf verwandte Passagen. Sie springen übrigens von einer Passage zur anderen einfach durch Anklicken des links zum Querverweis. Wollen Sie anschließend zurückspringen, nutzen Sie die Kombination der Tasten „Alt“ und „←“. Meine Webseite finden Sie hier zu Veranstaltungsplanung, weiterführenden Veranstaltungen im Master und allgemeinen Informationen. Meine Sprechstunde: n.V., aber am besten einfach direkt nach der Vorlesung. Herzlichen Dank an Wolfgang Dauth für die Genehmigung des Abdrucks von Abb. (9.3) sowie Clara Leppelmann für kritische Hinweise und Kommentare.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Orientierung .....	1
1.2	Regionalökonomie .....	3
1.3	Silicon Valley .....	7
1.4	„Omni-Channel“ .....	17
1.5	Kürzeste Wege (Dijkstra-Algorithmus) .....	20
<b>2</b>	<b>Neoklassische Regionalökonomie</b>	<b>23</b>
2.1	Einfachstmögliches interregionales Gleichgewicht .....	25
2.2	Grenzprodukt .....	28
2.3	Wohlfahrt .....	34
<b>3</b>	<b>Regionale Arbeitslosigkeit</b>	<b>41</b>
3.1	Erwartete Lohnsätze .....	43
3.2	Wanderungsgleichgewicht .....	45
3.3	Regionale Lohnsubventionen .....	46
<b>4</b>	<b>Keynesianische Regionalökonomie</b>	<b>53</b>
4.1	Exportbasistheorie .....	53
4.2	Koordinationsversagen .....	56
4.3	Rosenstein-Rodans <i>Big Push</i> .....	66
<b>5</b>	<b>Regionale Input-Output-Analyse</b>	<b>73</b>
5.1	Führungssektor Eisenbahnen .....	73
5.2	Verflechtung und Zirkularität .....	75
5.3	Multiplikatoreffekte .....	78

---

<b>6</b>	<b>Polarisation</b>	<b>83</b>
6.1	Gleichgewichtsloki .....	85
6.2	Multiple Gleichgewichte .....	88
6.3	Pfadabhängigkeit .....	90
<b>7</b>	<b>Region und Epidemie</b>	<b>99</b>
7.1	SIR-Modell .....	103
7.2	Epidemiekurve .....	111
7.3	Epidemiopolitik .....	114
<b>8</b>	<b>Evidenzbasierte Politik</b>	<b>119</b>
8.1	Wirksamer Impfstoff? .....	119
8.2	Evidenzbasierte Evaluation .....	125
8.3	Ausblick .....	131
<b>9</b>	<b>Empirie vs. Theorie</b>	<b>135</b>
9.1	Lohnkorrektur .....	135
9.2	Lohnstreuung .....	140
<b>10</b>	<b>Literatur</b>	<b>145</b>



## 7 Region und Epidemie

### Be. 110: Epidemie: Entscheidungen und Ressourcen

Ab spätestens Ende 2019 zog das Coronavirus um die Welt. Die bekannte Webseite der Johns Hopkins University versammelt aktuelle globale Fakten zur Epidemie. Die bekannte Webseite des Robert-Koch-Institut bietet eine Fülle von Informationen zum Verlauf der Epidemie in Deutschland. Das *New England Journal of Medicine* versammelt wissenschaftliche Papiere (ohne *paywall*) aus der Medizin und Epidemiologie, die Webseite des NBER, Beiträge aus der Ökonomie. (Diese Literatur erörtert u.a. zunehmend auch mittelfristige gesundheitliche Folgen der Infektion unter Genesenen.) Denn natürlich ist diese Epidemie nicht nur ein medizinisch-biologisches, sondern auch ein ökonomisches Thema – nicht so sehr, weil die Epidemie das wirtschaftliche Wachstum beschädigt, sondern weil die (Gesundheits-) Ökonomie sich zu Fragen zu optimalen Entscheidungen hinsichtlich der Allokation knapper Ressourcen stellt. Und wieder gilt es die optimale Allokation knapper Ressourcen zu prüfen: Krankenhäuser, Betten, Betten auf Intensivstationen, Beatmungsgeräte, Heilmittel, Telefonisten in Gesundheitsämtern und sich Risiken für die eigene Gesundheit aussetzendes Pflegepersonal.

### Be. 111: Epidemie und Region

Die Verbreitung von Covid-19 hält sich nicht an politische Grenzen; das Virus ist nicht nur interregional, sondern sogar international mobil, weil seine Träger es sind. Hat es trotzdem Sinn, hier Regionen und Länder zu unterscheiden? (1) Infektionen sind „räumlich korreliert“: Die Karte der Neuinfektionen im November 2021 aus Abb. (7.9) zeigt, daß Landkreise unterschiedlicher Farben sich nicht beliebig mischen. Vielmehr finden wir Landkreise einer gegebenen Farbe typischerweise neben Kreisen der gleichen Farbe. Haushalte sind mobiler zwischen benachbarten Kreisen als zwischen nicht benachbarten Kreisen. (2) Administrative Teileinheiten ergreifen unterschiedliche Maßnahmen: die Modalitäten des *testing*, *social distancing*, die Verpflichtung zum Maskentragen, das Ausmaß des *contact tracing*, ein *lockdown*, ein *Beherbungsverbot* u.ä. unterscheiden sich teils sogar von Landkreis zu Landkreis. (3) Medizinische Ressourcen variieren von Teileinheit zu Teileinheit und so auch die Fähigkeiten, Infektionsketten durch Heilung und Isolation zu unterdrücken. (4) Wichtige strukturelle Merkmale unterscheiden sich von Region zu Region. Manche Regionen sind dichter besiedelt,



(a) Pavillons Urban-Krankenhaus (Berlin)



(b) Hansaviertel (IBA Berlin)



(c) Sitzbank für Tbc-Kranke (Dublin/Bray)



(d) Vormalige Kuh-/Pferdeställe (Berlin)

Abbildung 7.1: Architekturen des *social distancing*

haben im Durchschnitt ältere Bewohner, sind eher Transitraum für internationalen Reiseverkehr oder Tourismus als andere, u.v.a.m. Nicht zuletzt hat Covid dauerhafte regionalökonomische Konsequenzen, weil es *work-at-home* hoffähig gemacht hat. In vielen Ländern wurden die Anziehungskräfte der Stadtzentren damit dramatisch geschwächt.

### Be. 112: Regionenvergleich als Experimentierfeld

Ein Blick auf die täglich aktualisierten Karten zeigt, wie sehr sich das empirische Infektionsgeschehen international, aber auch interregional unterscheidet. Das schließt nicht aus, daß dieser Blick nur eine Momentaufnahme ist. Wird Covid-19 endemisch, könnten sich Regionen auch wieder einander – im schlechten – angleichen. Ein Grund für die Prominenz regionalökonomischer Betrachtung ist die Möglichkeit, die unterschiedlichen Epidemie-Verläufe der Regionen mit den jeweils dort umgesetzten Politikmaßnahmen in Verbindung zu bringen – und hieraus für alle jeweils anderen Regionen zu *lernen* (Be. (5)). Mitze et al. (2020) etwa vergleichen die Erfahrungen in Jena, das das Maskentragen im öffentlichen

Nahverkehr und beim Einkaufen sehr viel früher zur Pflicht erklärte, mit denen anderer Städte bzw. Länder, die dies erst später taten. Diesen Autoren zufolge reduzierte das Gesichtsmaskentragen die Anzahl bekannter Covid-19-Fälle um zwischen „2.3% und 13%“. Glaeser et al. (2020) etwa streichen den Erfolg von *lockdowns* in kritischen Phasen der Epidemie konkret in New York heraus.

### Be. 113: Infektionskrankheiten und Immobilien

Covid-19 ist nur eine weitere in einer langen Reihe von bekannten Infektionskrankheiten. Zu ihnen zählen auch die Pest, die Cholera, die Masern, die Diphtherie, die Pocken oder die Tuberkulose. Deaton (2013, Kap. 2) bietet eine wirtschafts- und gesundheitsökonomische Analyse zu der Bedeutung solcher Infektionskrankheiten in der Vergangenheit. Die Krankenhäuser des späten 19.-ten Jahrhunderts setzen erste Erkenntnisse über Infektionen um. Räume und Fenster werden größer, um besser durchlüften zu können; Raumdecken werden höher. Das Krankenhaus am Urban in Berlin-Kreuzberg, das Krankenhaus Friedrichshain, die Tuberkulose-Heilstätten in Berlin-Buch oder im brandenburgischen Beelitz bestehen nicht aus einem einzigen monolithischen Gebäude, sondern teils aus vielen kleineren Bettenhäusern oder *Pavillons*, die durch Nebengebäude miteinander verbunden sind. Einige dieser heute noch erhaltenen, wenn jetzt auch anders genutzten Pavillons zeigt Abb. (7.1a), und die gerade beschriebenen Eigenschaften der Räume lassen sich heute noch gut an Fensterhöhe und -größe ablesen. Abb. (7.1b) zeigt ein Haus des Hansaviertels der Internationalen Bauausstellung (IBA) der 50er Jahre, in Berlin. In der Architektur der Moderne gebaut will dieses Haus der Forderung nach „Licht und Luft“ nachkommen, die den Erfahrungen leichter Übertragbarkeit von Krankheiten in eng bewohnten Hinterhöfen entspringt. Abb. (7.1c) zeigt eine Sitzbank, die *social distancing* an der frischen Luft in Zeiten der Tuberkulose ermöglichte. Abb. (7.1) zeigt Kuh- oder Pferdeställe des 19. Jahrhunderts im höchst dichtbiedelten Berlin-Kreuzberg, die immerhin in einem separaten Gebäude untergebracht sind.

### Be. 114: Covid-19, Regionalökonomie und Immobilienwirtschaft

Liu/Su (2021) zeigen die immobilienwirtschaftlichen und regionalökonomischen Konsequenzen der Epidemie für amerikanische Städte. Viele Autoren haben im letzten Jahr vermutet, daß die Corona-Krise dicht besiedelten Regionen zum Nachteil gereicht. Abb. (7.2) ist die erste empirische Analyse (hier: aus den USA) zu dieser Frage, und sie bestätigt die Erwartung dieser Beobachter. Auf der linken Seite der Abbildungen finden wir in drei Diagrammen, daß der Bestand leerstehender Häuser (*inventory*) am Stadtrand stärker gefallen ist als im Zentrum, daß Immobilienpreise am Stadtrand stärker gestiegen sind als im Zentrum und daß Mieten im Stadtzentrum nicht so stark gestiegen wie die am Rand. Darüber hinaus zeigt die rechte Seite von Abb. (7.2), daß die 25 größten Städte der USA sich relativ zu allen anderen Städten der USA weniger dynamisch

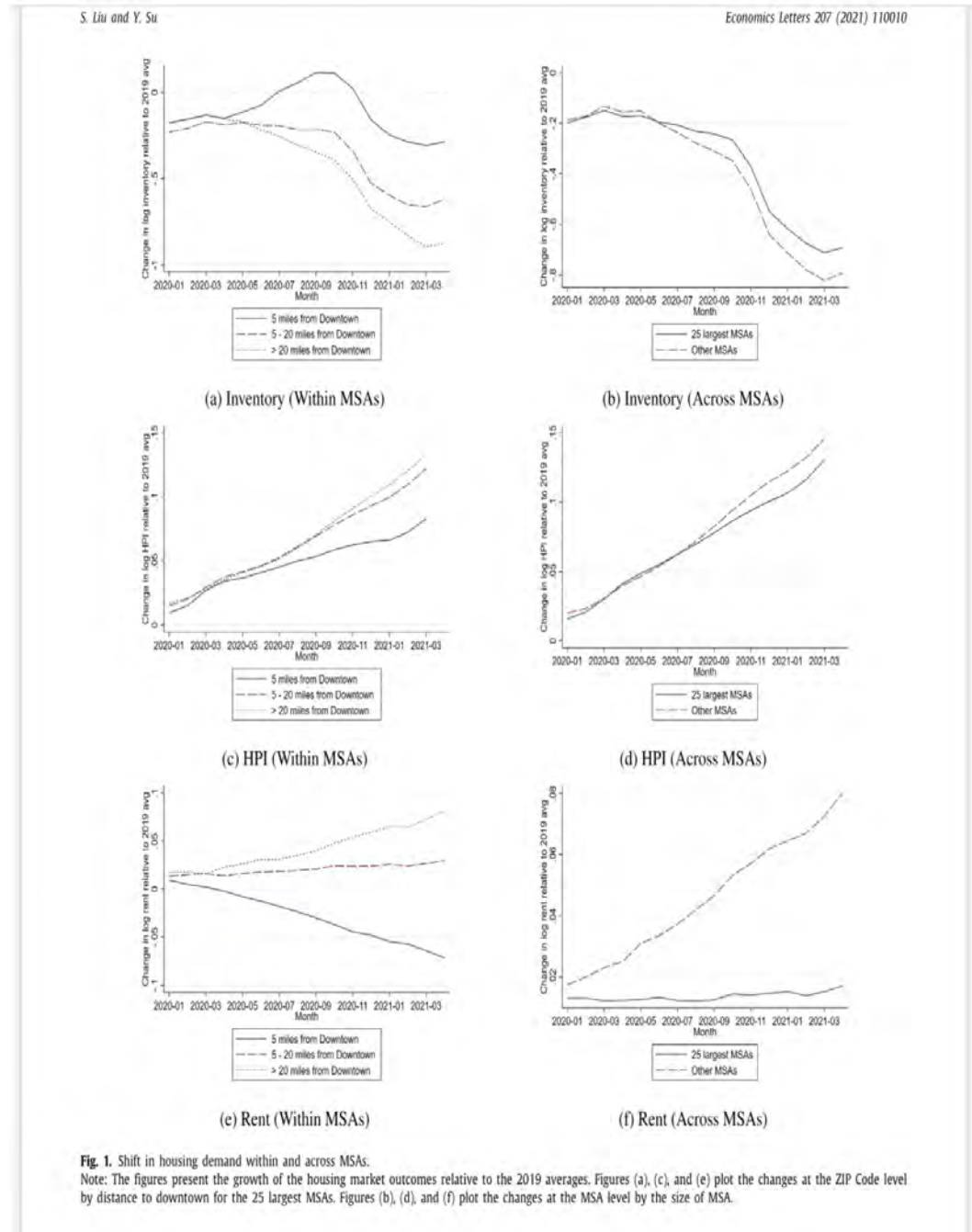


Abbildung 7.2: USA: Wanderungen in die Vorstädte (Liu/Su (2021))



entwickeln. *Das Corona-Virus justiert die Balance zwischen dicht- und weniger dicht besiedelten Regionen neu: Haushalte meiden dicht besiedelte Räume zusehends.* Überlegen Sie, welche *drastischen Signale* von diesen Preisentwicklungen für Entwickler, für die Immobilienwirtschaft ausgehen. Der zukünftige Neubau der USA wird sich aus den Verdichtungsräumen wegorientieren. Rekapitulieren wir zum Schluß den regionalökonomischen Bezug der Corona-Krise, in aufsteigender Relevanz: (1) Regionen sind ein „Laboratorium“ (vierter Punkt in Be. (5)). Vergleiche von Regionen informieren gesundheitspolitische Entscheidungen. (2) Methoden der Epidemiologen bereichern die Methoden der Ökonomen: Evidenzbasierte Politik greift auf die Vorteile des *difference-in-difference*-Ansatzes zurück (Be. (138)). Epidemiologie informiert Regionalökonomie. (3) Verdichtete Regionen werden relativ weniger attraktiv.

### Be. 115: Covid-19 und Städtebau

Der sehr sehenswerte, für Studierende der Immobilienwirtschaft fast schon zum Pflichtprogramm zählende Film „Citizen Jane: Battle for the City“ rekapituliert die große städtebauliche Kontroverse des 20. Jahrhunderts am Beispiel New Yorks. In dieser Kontroverse spielen epidemiologische Argumente (mindestens vordergründig) eine wichtige Rolle. Auf der einen Seite steht dort Robert Moses, der als Planer der Stadt hochverdichtete Quartiere abreißt und durch die Architektur der Moderne (vulgo: Wohntürme plus Abstandsflächen) ersetzt. Zur Begründung verweist er auf die in den alten Vierteln erleichterte Verbreitung von „diseases“ (Infektionskrankheiten). (Nicht auszuschließen ist allerdings, daß andere Motive ebenfalls eine Rolle gespielt haben können. *Slum Clearance* stand u.a. auch für ethnische Diskriminierung.) Auf der anderen Seite steht Jane Jacobs, die sich zugunsten der Vorteile hochverdichteter Viertel ausspricht. Diese städtebauliche Kontroverse spiegelt sich anschaulich auch in der Nachkriegsentwicklung deutscher Städte. – Le Corbusier ist international der zentrale Vertreter einer Entdichtung. Seine Häuser stehen u.a. auf Stelzen, Arbeiten und Wohnen sind räumlich getrennt. In seiner *Charta von Athen* fordert er „Licht und Luft“ für alle.

## 7.1 SIR-Modell

### Be. 116: Das SIR-Modell

Das berühmte SIR-Modell der Epidemiologie von Kermick/McKendray (1927) erklärt uns die zentrale Eigenschaft einer ungehinderten Epidemie: ihr explosives Wachstum am Anfang, ihr anschließendes ähnlich schnelles Abebben, vgl. die Verläufe der „Großen Pest“ in London 1665-1666 und einen Masternausbruch in New York in Abb. (7.3). Dieses Modell ist zwar ein frühes Modell der Epidemiologie, einer Teildisziplin der Theoretischen Biologie, aber es ähnelt gleichzeitig

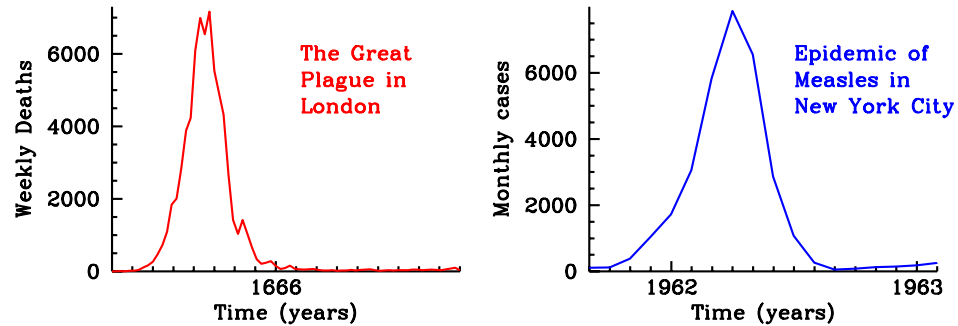


Abbildung 7.3: „Große Pest“ London, Masern New York (Earn (2008))

makroökonomischen Modellen. Es hat: Aggregate, Verhaltensannahmen an diese Aggregate, Gleichgewicht und Dynamik, und es bietet Ansatzpunkte für das Eingreifen der Politik, die auf gesellschaftliche Wohlfahrt zielt. Die folgende Darstellung lehnt sich an Heuser (1986), Earn (2008), Akira-Toda (2020) und Avery et al. (2020) an. – Als Ausblick auf die Themen dieses Kapitels vergleichen wir die historischen Pest- und Masern-Epidemien aus Abb. (7.3) mit dem bisherigen Verlauf der Corona-Epidemie Stand November 2021 (Abb. (7.4)). Beiden Abbildungen ist zu Epidemiebeginn der exponentielle Verlauf gemeinsam. Allerdings ebbt die Masern und die Pest nach diesem Ausbruch auch schnell wieder ab. Dieses Abebben sehen wir für die Neuinfektionen durch das Coronavirus zwar auch; aber ihm folgt unweigerlich auch wieder der nächste Ausbruch. In diesem Unterschied spiegelt sich die heute bessere Kombination aus wissenschaftlichem Verständnis und Politik: *lockdowns* u.ä. können die vollständige Ausbreitung des Virus temporär verhindern. (Wird der *lockdown* aber wieder aufgehoben, flackert die Epidemie allerdings wieder auf.) Wir erhalten den „wellenförmigen“ Verlauf aus Abb. (7.4). – Die folgenden Ausführungen stellen mit SIR das einfachste Modell vor; insbesondere ignorieren wir vorerst die seit Anfang 2020 bestehende Möglichkeit der *Impfung*.

**Be. 117:** *Susceptibles, Infecteds, Removeds*

Wir unterscheiden drei Gruppen (Teile, Aggregate) der Gesamtbevölkerung  $N$ . Die *susceptibles*,  $S$ , sind all die, die noch nie infiziert waren – und daher noch infiziert werden können. Zu Beginn der Epidemie sind alle Regionsbewohner *susceptible* (anfällig). Die *infecteds*,  $I$ , sind diejenigen, die sich mit dem Virus infiziert haben und aber noch nicht geheilt – oder verstorben – sind. Sie sind insbesondere auch *infektiös*. Sie können das Virus an diejenigen, die noch nicht infiziert, immun oder verstorben sind, weiterreichen. Schließlich gibt es noch die *removeds*,  $R$ . Das sind die, die das Infektionsgeschehen nicht mehr interessiert

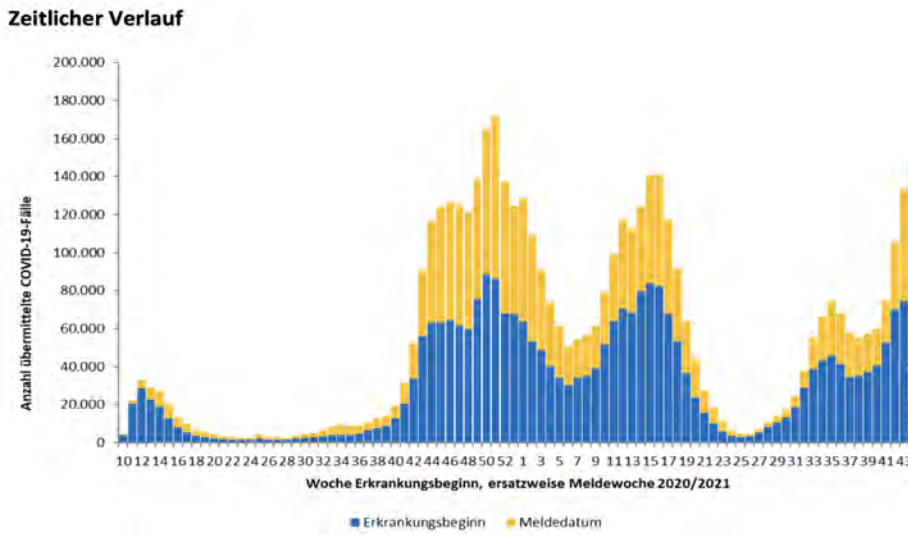


Abbildung 7.4: Coronavirus in Deutschland (Neuinfektionen lt. RKI, 12.11.2021)

– entweder weil sie immun oder weil sie verstorben sind. Zusammengefaßt: Es gibt die *susceptibles* (Anfälligen), die *infecteds* (Infektiösen) und die *removeds* (Resistenten) – insgesamt „SIR“.

#### Be. 118: *Susceptibles*, Infizierbare

Das SIR-Modell – und uns – interessiert die Entwicklung von  $S$ ,  $I$  und  $R$  in der Zeit, erfaßt durch den Zeitindex  $t$ . Es interessiert sich also für die Funktionen  $S(t)$ ,  $I(t)$  und  $R(t)$ . Ausgangspunkt des Modells ist, daß die Anzahl der *susceptibles* im Zeitablauf fällt, daß also die Änderungsrate  $dS/dt \equiv \dot{S}$  negativ ist. Hinter diesem Ausgangspunkt steht die Annahme, daß am Anfang zwar alle *susceptible* sind, später dies aber nur noch der Teil der Bevölkerung ist, der noch nicht infiziert worden (oder verstorben) ist. Bereits Infizierte sind jetzt nicht mehr *susceptible*, d.h. immun.<sup>1</sup> Immer weniger Individuen stehen dem Virus im Zeitablauf als „Wirt“ zur Verfügung. Noch etwas konkreter unterstellt das Modell:

$$\dot{S} = -\beta S(t)I(t)/N < 0. \quad (7.1)$$

Schauen wir uns den Ausdruck auf der rechten Seite an, den Zuwachs an Infizierten an *einem* Tag, die Neu-Infizierten eines Tages.<sup>2</sup> Die täglichen Neuinfektionen – und gleichzeitig die tägliche Reduktion an *susceptibles*, Infizierbaren – ist hier

<sup>1</sup>Für Covid-19 ist diese Annahme nicht richtig. Auch Genesene können sich erneut anstecken.

<sup>2</sup>Die Wahl anderer Zeiteinheiten ist ebenfalls möglich, z.B. eine Woche.

dem Produkt aus *susceptibles* und *infecteds*,  $SI$  proportional. Warum? Jeder Infizierte (Be. (119)) trifft mit Wahrscheinlichkeit  $S(t)/N$  auf einen *susceptible*. (Nicht jeder soziale Kontakt ist infizierbar. Die bereits Infizierten sind es ja nicht mehr, und auch nicht die *removeds*.) Die Wahrscheinlichkeit, diesen dann auch anzustecken, sei  $\beta$ ,  $0 < \beta < 1$  (das ist die sog. *Infektionsrate*). Damit infiziert jeder Infektiöse „erfolgreich“ seinen Kontakt mit Wahrscheinlichkeit  $\beta S(t)/N$ . Bei insgesamt  $I(t)$  Infizierten resultiert schließlich für Tag  $t$  der Ausdruck auf der rechten Seite von Gl. (7.1).

**Be. 119: *Infecteds*, Infektiöse**

Von den täglichen Neuinfizierten  $\dot{S}$  ganz klar zu unterscheiden ist die tägliche Veränderung im Bestand der Infizierten,  $\dot{I}$ . Denn dort gibt es ja nicht nur Zugänge (die Neuinfizierten eben), sondern auch Abgänge (die immun Gewordenen und die Verstorbenen),  $\dot{R}$  (Be. (120)). Es ist also

$$\dot{I} = \beta(S(t)I(t))/N - \gamma I(t). \quad (7.2)$$

Den ersten Term auf der rechten Seite erkennen wir als die täglichen Neuinfektionen wieder. Der zweite Term entspricht den täglichen *removeds*,  $\dot{R}$ , mit  $1 > \gamma > 0$  als sog. *Ausfallrate*. Beachten Sie jetzt: Die beiden Gleichungen (7.1) und (7.2) sind nicht nur interdependent: Sie sind simultan zu lösen. Sie sind auch Differentialgleichungen: Sie verknüpfen die Niveaus der beiden zentralen Variablen  $S$  und  $I$  mit ihren Veränderungen. Diese beiden Gleichungen zu lösen (wir wollen es nicht), hieße, Funktionen  $S(t)$  und  $I(t)$  zu finden, die dieses Zwei-Gleichungssystem zu jedem Zeitpunkt  $t$  erfüllen. Besonders interessant wäre dabei die Funktion  $I(t)$ . Sie gibt an, wie sich die Anzahl der Infizierten im Zeitablauf verändert. Sie bildet das tatsächliche Infektionsgeschehen im Zeitablauf ab.

**Be. 120: *Removeds*, Resistente**

Die dritte Gleichung des SIR-Systems ist

$$\dot{R} = \gamma I(t). \quad (7.3)$$

Sie beschreibt, wie sich die Anzahl der *removeds* täglich verändert. Sie ist das Produkt der sog. *Ausfallrate* (mißverständlicher Begriff) mit der Anzahl der Infizierten in  $t$ . Wie beschrieben fassen wir hier die Verstorbenen und die jetzt immunen Überlebenden zusammen. Weder die eine noch die andere Gruppe ist (im Modell, das gilt nicht für Covid-19) weiter *susceptible*; und weder die eine noch die andere Gruppe ist *infected* i.S.v. infektiös. Die Literatur spricht hier auch von einem Modell mit drei *compartments* oder „Abteilen“. Jedes Gesellschaftsmitglied befindet sich anfangs im Abteil der *susceptibles*. Je nach Epidemieverlauf durchwandert es anschließend (und nicht fröhlich, d.h. zu oft erheblichen Leiden und also gesellschaftlichen Kosten) ggf. auch noch die zwei anderen Abteile. – Im Aggregat gilt nach Aufaddieren aller drei Gleichungen:  $\dot{S} + \dot{I} + \dot{R} = 0$ .

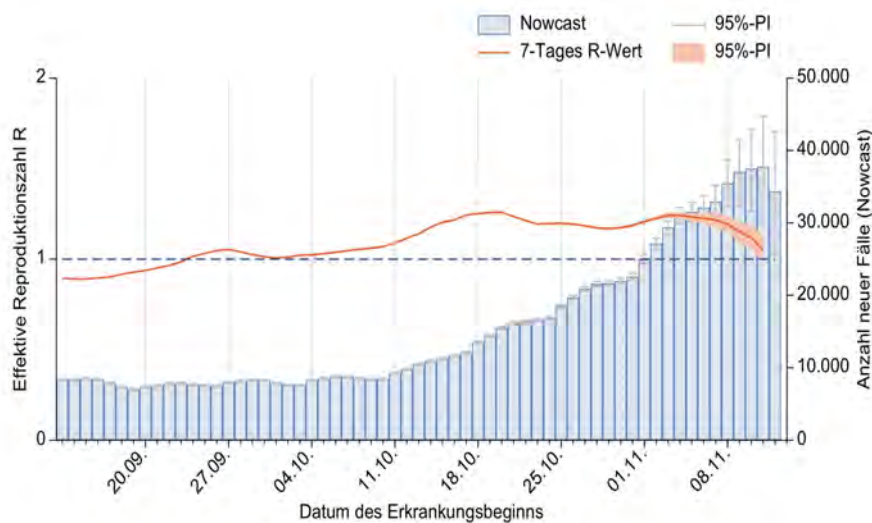


Abbildung 7.5: Effektive Reproduktionszahl (RKI, 16.11.2021)

**Be. 121: Wieviele Infizierbare steckt ein Infektiöser an?**

Ein sehr einfach zu bestimmender, aber gleichzeitig hochdramatischer Moment innerhalb des SIR-Modells ist dort, wo sich entscheidet, ob die tägliche Anzahl der Neu-Infektionen steigt ( $\dot{I} > 0$ ) oder fällt ( $\dot{I} < 0$ ). Betrachten wir mit unseren bisherigen Konzepten jetzt das Produkt

$$(1/\gamma) \cdot (S(t)/N) \cdot \beta \quad \text{oder} \quad (\beta/\gamma) \cdot (S(t)/N). \quad (7.4)$$

Der erste Faktor,  $1/\gamma$ , links erfaßt die Tage, die ein Infizierter durchschnittlich infektiös ist. Das Produkt der anderen beiden Faktoren links erfaßt die Neu-Infektionen, die jeder Infizierte je Tag auslöst (Be. (118)). Insgesamt beschreibt das obenstehende Produkt (ob links oder rechts) also die Anzahl Neu-Infektionen, die jeder Infizierte *während seiner gesamten infektiösen Phase* auslöst. Diese Anzahl heißt *effektive Reproduktionszahl*  $r_t$  und bezieht sich auf den Tag  $t$ . Sie ändert sich sicherlich im Epidemieverlauf. Befinden wir uns am Beginn einer Epidemie (also zu  $t = 0$ ), dann gilt in etwa  $S(0) = N$ . Die effektive Reproduktionszahl vereinfacht sich dann zu

$$(1/\gamma) \cdot \beta \quad \text{oder} \quad \beta/\gamma \quad (7.5)$$

Dieser einfache Quotient ist die sog. *Basis-Reproduktionszahl*,  $r_0$ . Sie beschreibt die Anzahl der *susceptibles*, die „der erste Infizierte“ ansteckt, *wenn also die Epidemie noch gar nicht richtig begonnen hat*. Zu Beginn der Epidemie dürfen wir uns an  $r_0$  orientieren; später müssen wir  $r_t$  heranziehen. Kraft ihrer Interpretation ist intuitiv klar: Liegt die Reproduktionszahl über 1, dann nimmt die Anzahl

Maßnahme	Wirkung ...
Abstand	$\beta \downarrow$
Hygiene	$\beta \downarrow$
Maskenpflicht	$\beta \downarrow$
Testen ...	$\beta \downarrow$
Reiseverbote u.ä.	$\beta \downarrow$
<i>lockdown</i>	$\beta \downarrow$
Quarantäne, Isolation	$\gamma \uparrow$
Innovative Behandlung	$\gamma \uparrow$
Resignation?	$S(t) \downarrow$

Tabelle 7.1: Coronapolitik u. Wirkungen auf die effektive Reproduktionszahl

der Neuinfizierten täglich weiter zu. Liegt sie unter eins, nimmt sie täglich weiter ab. Insofern ist die Zahl 1 ein zentraler *Schwellenwert*. Im Vergleich mit 1 entscheidet sich, ob die Infektionskette „abstirbt“ oder „explodiert“ – zwei sehr gegensätzliche Konsequenzen, die aber anfänglich sehr nahe beieinander liegen und die Dringlichkeit schnellen politischen Handelns motivieren (Be. (123)). In diesen Beobachtungen nimmt die Beschäftigung der Öffentlichkeit mit dem täglichen Vergleich der (geschätzten aktuellen) effektiven Reproduktionszahl mit 1 ihren Ursprung (Abb. (7.5)). Mitte November 2021 liegt  $r_t$  bei 1.05. Die Zahl der Infektösen  $I$  sollte also erst einmal weiter ansteigen.

### Be. 122: Basis- und effektive Reproduktionszahl

Versuchen wir einen alternativen, vielleicht noch einfacheren Zugang zur Reproduktionszahl. Schauen wir zuerst einfach einmal auf die Funktion  $\dot{I}$  in Gl. (7.2). Ausklammern von  $I$  auf der rechten Seite führt zu

$$\dot{I} = (\beta S(t)/N - \gamma) I(t) \quad (7.6)$$

Halten wir zuerst fest, daß in den ersten Tagen nach Auftreten des ersten Infizierten, also für  $t > 0$ ,  $I(t)$  zwar sehr klein, aber trotzdem schon strikt positiv ist. Dann gilt aber auch: Die Änderungsrate der Neu-Infektionen,  $\dot{I}$ , ist lt. Gl. (7.6) positiv (negativ, Null) genau dann, wenn die Differenz in Klammern auf der rechten Seite von (7.6), also  $(\beta S(t)/N - \gamma)$ , positiv (negativ, Null) ist. Äquivalent:

$$\dot{I} \geq 0 \quad \Leftrightarrow \quad (\beta/\gamma) \cdot (S(t)/N) \geq 1. \quad (7.7)$$

Der Ausdruck auf der linken Seite der zweiten Ungleichung ist die schon bekannte *effektive Reproduktionszahl*. Deren (vom RKI geschätzten) zeitlichen Verlauf

sowie (rot eingezeichnet) den Schwellenwert 1 zeigt die Graphik in Abb. (7.5).  
 – Unsere Überlegungen lassen sich noch weiter vereinfachen, wenn wir erst am Beginn einer möglichen Epidemie stehen. Denn dann gilt ja gerade  $S(0) = N$ , so daß  $N/S(0) = 1$ . Zu Beginn einer möglichen Epidemie gilt also sogar:

$$\dot{I} \geq 0 \quad \Leftrightarrow \quad \beta/\gamma \geq 1. \quad (7.8)$$

Hier sehen wir wieder den bereits bekannten Vergleich der bekannten *Basis-Reproduktionszahl* mit der Zahl 1. – In jedem Fall gilt: Liegt die (relevante) Reproduktionszahl über 1, kommt es zu einem Anstieg der Zahl der *Neu-Infizierten*. Liegt sie darunter (und nur dann), kommt es zu einem Rückgang.

### Be. 123: Hebel der Coronapolitik

Politische Maßnahmen können die wichtige Reproduktionszahl beeinflussen. Wir unterscheiden solche Maßnahmen, die  $\beta$  senken, von solchen, die  $\gamma$  steigern. (Gar nichts zu tun und darauf zu warten, daß die Reproduktionszahl allein deswegen fällt, weil die Anzahl der *susceptibles* mit zunehmenden Infektionen immer weiter zurückgeht, ist „passive Politik“. In unterschiedlichen Phasen diskutierten Schweden und Großbritannien diese Strategie.) Bekannte Maßnahmen, die soziale Interaktionen reduzieren, sind *Maskentragen*, *Abstandhalten*, *home office*, *stay-at-home-orders* oder sogar ein richtiger *lockdown*. Sie alle helfen  $\beta$  zu reduzieren. (Covid-19 hat eine zusätzliche bekannte Komplikation, die in unserem einfachen Modell noch nicht richtig berücksichtigt ist: Träger des Virus haben u.U. keine Symptome. Umso wichtiger wird in diesem Kontakt das *Testen*, die *Kontaktnachverfolgung* sowie *Isolation und Quarantäne* der Infizierten. Auch diese Maßnahmen lassen sich grob dem erwünschten Rückgang von  $\beta$  zuordnen.) Am besten ist natürlich ein Impfstoff (Vakzin) (Be. (132)). Manche Medikamente mögen gleichzeitig zu einem Anstieg von  $\gamma$  führen, wenn sie die Viruslast bzw. die Dauer der infektiösen Phase reduzieren. – Unabhängig von den gut einschätzbaren qualitativen Effekten (vgl. Tabelle (7.1)) drängt sich die Frage nach den quantitativen Effekten auf, i.e. wie stark sind diese Effekte, und was kosten sie jeweils?

### Be. 124: Exponentielles Wachstum

Klären wir jetzt die (anhand der Verläufe aus Abb. (7.3) motivierte) Befürchtung eines exponentiellen Wachstums der Neu-Infektionen. Am Anfang der Epidemie vereinfacht sich Gl. (7.2)

$$\dot{I} = (\beta - \gamma) I(t). \quad (7.9)$$

Diese Gleichung ist immer noch eine Differentialgleichung – aber eine einfache (die einfachstmögliche). In ihr ist die Veränderung der Infektiösen in der betrachteten Zeitperiode dem Niveau der Infektiösen *proportional*. Nach Umformen in

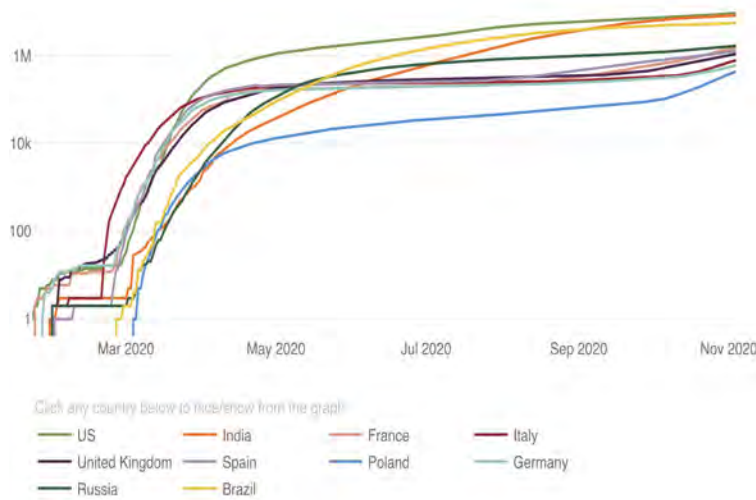


Abbildung 7.6: Logarithmierter Graph der Neu-Positiv-Getesteten (JHU)

$\dot{I}/I = \beta - \gamma$  könnte man auch sagen: Die Wachstumsrate ist einfach gleich  $\beta - \gamma$ , also kurz- und mittelfristig konstant. Eine *positive* konstante Wachstumsrate der Neuinfektionen ist dramatisch: Z.B. sind +2% am Anfang, bei insgesamt 100 Infektiösen, nicht schlimm: es kommen ja nur zwei dazu. Aber +2% Neuinfektionen, wenn bereits 10 Mio. infiziert sind, ist schlimm: Dann kommen in der gleichen Zeitperiode 200,000 Neuinfizierte (!) hinzu. Unser Gesundheitssystem könnte mit diesen Zahlen nicht umgehen; in diesem Kontext sind die Warnungen oder Befürchtungen der Epidemiologen (z.B. Drosten, Lauterbach, u.a.) zu sehen. – Formal ist die Lösung der Differentialgleichung in (7.9) gerade

$$I(t) = e^{(\beta - \gamma)t}, \quad (7.10)$$

und diese Lösungsfunktion ist tatsächlich die Exponentialfunktion. Deren Exponent ist übrigens nicht nur  $(\beta - \gamma)t$ , sondern alternativ auch  $(r_0 - 1)\gamma t$ . Die Wachstumsrate der Infektiösen Entwicklung ist proportional dem Überschuss der Basis-Reproduktionszahl gegenüber Eins.

#### Be. 125: Logarithmische Darstellung; *bending the curve*

Wenn wir beide Seiten von Gl. (7.10) logarithmieren, dann erhalten wir

$$\ln I(t) = (\beta - \gamma)t$$

Der Logarithmus der Neu-Infektionen bzw. Infektiösen ist also einfach *linear* in der Zeit. Viele Graphiken, die anfangs in den Medien zu finden waren, trugen die *logarithmierten* Covid-19-Getesten (was natürlich nicht das gleiche war wie



die tatsächlich Neuinfizierten, es unterschätzte sie vermutlich dramatisch) gegen die Zeit ab. In Abb. (7.6) hat die Ordinate einen logarithmischen Maßstab. Deshalb ist liegt die Referenzlinie für 100 positive Getestete auch nur halb so hoch wie die für 10,000 ( $\ln(100) \approx 0.5 \ln(10000)$ ).<sup>3</sup>) Die Grundidee der logarithmischen Darstellung ist die Hoffnung auf schnelle Lesbarkeit: Ist der gezeichnete empirische Graph linear („eine Gerade“), wirft das sofort den Verdacht auf, hier handele es sich um einen exponentiellen Wachstumsprozeß. Tatsächlich verliefen viele dieser Graphen in den ersten Tagen und Wochen linear; später, als (und weil) dann schon die ersten Politik-Maßnahmen griffen, „bogen sie sich herunter“. Deshalb heißen diese Maßnahmen auch *bending the curve*.

## 7.2 Epidemiekurve

### Be. 126: Formale Lösung?

Die folgende grobe Fassung einer Lösung des Systems aus Gl. (7.1) und (7.2) nutzt die Annäherungen  $\dot{S} \approx \Delta S/\Delta t$  und  $\dot{I} \approx \Delta I/\Delta t$  aus.<sup>4</sup> Dann dividiert sie Gl. (7.2) durch Gl. (7.1). Hieraus folgt mit Hilfe von  $\Delta I/\Delta S \approx dI/dS$  approximativ

$$dI/dS = -1 + \frac{\gamma N}{\beta S} = -1 + \frac{\rho}{S} \quad (7.11)$$

In der zweiten Gleichung setzten wir zur Abkürzung  $\rho = \gamma N/\beta$ . In Worten: Wir erhalten einen Ausdruck, der die Abhängigkeit von  $I$  gegenüber  $S$  präzisiert. Das ist jetzt nicht die eigentlich gesuchte Abhängigkeit von  $I$  nach der Zeit  $t$ ! Aber es ist immerhin die Ableitung von  $I$  nach  $S$  – und dazu ein einfacher Ausdruck! Die Funktion  $I(S)$  finden wir anschließend durch Integration. (Unbestimmtes) Integrieren über der linken und rechten Seite dieser Gleichung gibt

$$I(S) = -S + \rho \ln S + C, \quad (7.12)$$

mit  $C$  als einer gleich zu bestimmenden, noch unbekanntem Konstante. (Überprüfen Sie diese Behauptung, indem Sie Gl. (7.12) nach  $S$  ableiten.)

### Be. 127: Epidemiekurve

Die Gleichung (7.12) beschreibt, wie wir gleich sehen werden, eine theoretische Epidemiekurve. *Ein* Punkt dieser Epidemiekurve ist schon bekannt. Denn am Anfang der Epidemie ist (fast exakt)  $S = N$  und  $I = 0$ . Also können wir schreiben:

$$0 = -N + \rho \ln N + C \quad (7.13)$$

<sup>3</sup>Genau genommen ist die Information an der Ordinate nicht richtig: Dort müßte  $\log(100)$ ,  $\log(10000)$ , etc. stehen.

<sup>4</sup>Die präzise Herleitung der Lösung findet sich in Heuser (1986, S. 316-317).

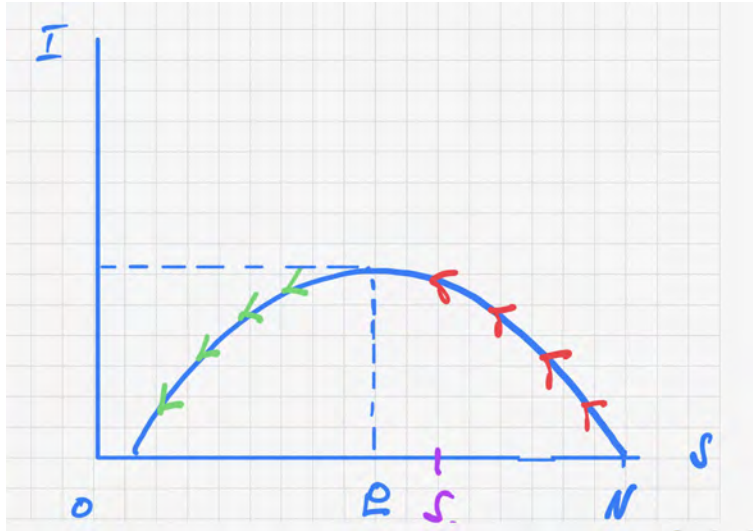


Abbildung 7.7: Theoretische Epidemiekurve ohne Gegenmaßnahmen

und damit also  $C = N - \rho \ln N$ . Jetzt haben wir die Konstante  $C$  bestimmt. Wir substituieren den für sie gefundenen Wert zurück in Gleichung (7.12) und erhalten schließlich das Ergebnis:

$$I(S) = (N - S) + \rho \ln \left( \frac{S}{N} \right). \quad (7.14)$$

Diese Gleichung beschreibt die *Trajektorie* im  $(S, I)$ -Diagramm, die die Gesellschaft durchläuft – kurz: die *Epidemiekurve* der Gesellschaft. Wie erwähnt beschreibt sie nicht die Entwicklung der Infektiösen in Abhängigkeit von der Zeit. Aber sie beschreibt die Entwicklung der Infektiösen in Abhängigkeit von den Anfälligen – und deren Entwicklung kennt nur eine Richtung (nach unten); sie ist damit wenigstens noch ein ungefähres Maß für die verstrichene Zeit.

### Be. 128: Moment der Krise

Jetzt bestimmen wir noch die Maximalstelle dieser Epidemiekurve. Ableiten von Gl. (7.14) nach  $S$  führt wieder zurück zu Gl. (7.11). Gleichnullsetzen dieser Gleichung und Umstellen nach  $S$  identifiziert dasjenige Niveau an *susceptibles*, zu dem die Anzahl der Infektiösen – damit auch die befürchtete Überlastung des Gesundheitssystems – maximal ist. Und das ist ausgerechnet

$$S^* = \rho.$$

Halten wir jetzt kurz noch einmal fest, was Gl. (7.14) nicht liefert: Sie liefert keinen Graphen der Infektiösen im Zeitablauf – etwa analog zum empirischen



Abbildung 7.8: Theoretische Epidemiekurve mit Gegenmaßnahmen

Graphen aus Abb. (7.3). (Einen solchen, ihm ähnlichen aussehenden Graphen liefern allerdings numerische Simulationen des SIR-Modells. Vgl. dazu die numerischen Simulationen bzw. Animationen unter [wikipedia/SIR-Modell](https://de.wikipedia.org/wiki/SIR-Modell).) Aber auch so können wir die zentralen Aussagen des SIR-Modells auf graphische Weise diskutieren.

### Be. 129: Epidemie – Keine Epidemie

Abb. (7.7) und (7.8) zeigen zwei mögliche theoretische Epidemiekurven. Schauen wir zuerst auf Abb. (7.7). Liegt  $\rho$  links des anfänglichen  $S$  (vgl.  $S_0$  lila), dann wird anschließend die Anzahl der Neuinfektionen (in Richtung der roten Pfeile) entlang der Epidemiekurve, von rechts nach links gelesen, *steigen*. Eine *Epidemie* mit einer täglich immer größer werdenden Anzahl Neuinfektionen bricht los. Erst wenn die Anzahl der *susceptibles* auf das Niveau von  $S = \rho$  gefallen ist, wird die Epidemie wieder abklingen. Eine schöne Aussicht ist das nicht, und es kann länger dauern, als das Diagramm (das ja nur eine Skizze ist) nahelegt. Schauen wir alternativ auf Abb. (7.8). Liegt  $\rho$  dagegen rechts des anfänglichen (lila)  $S_0$ , dann klingt die Anzahl der Neuinfektionen (in Richtung der grünen Pfeile) entlang der Epidemiekurve, von rechts nach links gelesen, sofort wieder ab. Es kommt schon noch zu neuen Infektionen. Aber diese Neuinfektionen werden täglich weniger. Etwas anders: Selbst im günstigen Szenario der Abb. (7.8) ist Covid-19 nicht notwendig verschwunden. Schwere Krankheitsverläufe, langanhaltende Symptome, Todesfälle sind jederzeit möglich – etwa wenn Zuwanderung Infizierter aus anderen Regionen stattfindet. „Günstig“ ist

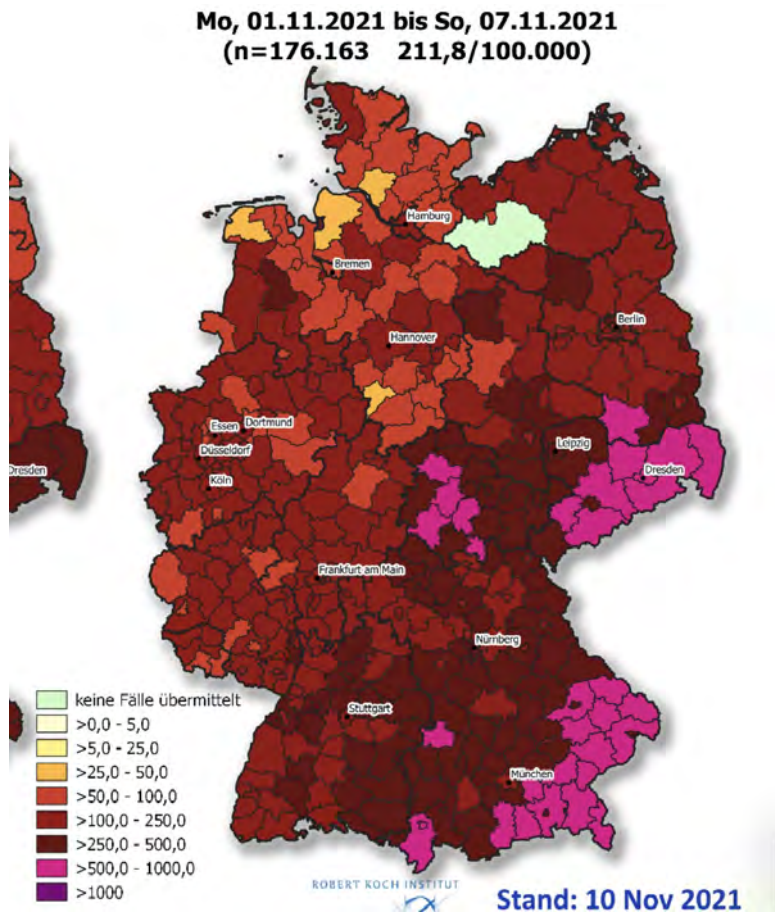


Abbildung 7.9: Deutschland-Karte der Neuinfektionen (10.11.2021), aus: RKI

hier lediglich, daß auf ein Aufflackern dieser Infektionskrankheit automatisches Wieder-Abklingen folgt.

### 7.3 Epidemiopolitik

#### Be. 130: Die Bedeutung politischer Maßnahmen

Welcher Fall, welche Abbildung relevant ist, hängt entscheidend davon ab, mit welcher Epidemiekurve wir rechnen müssen. Üb. (7.3) zeigt Ihnen, welchen entscheidenden Einfluß  $\rho$  – und damit im Hintergrund wieder  $\beta$  und  $\gamma$  bzw. die hinter ihnen stehenden Politikmaßnahmen der Tab. (7.1) – für die Lage der Epidemiekurve spielen. Abb. (7.8) zeigt nur deswegen einen günstigeren Verlauf

der Epidemiekurve, weil dort politische Maßnahmen der Eindämmung unterstellt sind. Liegen diese Voraussetzungen nicht vor, ist Abb. (7.7) relevant. Ein letzter Punkt (und ein gewisser, sehr schwacher Trost): Selbst im ungünstigsten Szenario werden sich nie alle anstecken. Die Epidemiekurve aus Abb. (7.7) schneidet die Abszisse nicht im Ursprung, sondern erst später. Die Intuition hierfür ist mit unseren Konzepten plausibel: Haben sich erst einmal hinreichend viele infiziert, bilden sie einen Schutzwall um die verbleibenden *susceptibles*. In ihnen findet das Virus keinen Träger, der sie auch noch zu jenen bringen könnte.

### Be. 131: Herdenimmunität

Herdenimmunität entsteht, wenn  $S$  auf den Schwellenwert  $\rho$  gefallen ist. Mit  $\rho$  ist die Anzahl der *susceptibles* so niedrig geworden, daß jeglicher lokale Ausbruch automatisch, von allein, in Ermangelung hinreichend vieler Infizierbarer, er stirbt. Wir definieren Herdenimmunität als den Bevölkerungsanteil

$$1 - \rho/N \quad \text{oder} \quad 1 - 1/r_0,$$

der *resistent* sein muß. Advokaten der Herdenimmunität hoffen auf einen schnellen Rückgang der *susceptibles* infolge einer schnellen Ausbreitung des Virus. Sie streben in Abb. (7.7) eine zügige *Bewegung auf der Epidemiekurve* an. Advokaten der oben beschriebenen Maßnahmen wie *social distancing*, *contact tracing* etc. dagegen streben die *Verschiebung der Epidemiekurve* von Abb. (7.7) nach Abb. (7.8) an. Zwei Zahlenbeispiele illustrieren die Interpretation des Schwellenwerts  $N - \rho$  als Herdenimmunität. Ist die Basis-Reproduktionszahl  $r_0 = 2$ , dann wird Herdenimmunität  $1 - 1/r_0$  erst erreicht, wenn die Hälfte der Bevölkerung *removed*, *resistent* (bzw. später ggf. geimpft) geworden ist. Ist die Basis-Reproduktionszahl  $r_0$  dagegen 3, dann erfordert Herdenimmunität einen *resistenten* Anteil von  $1 - 1/r_0 = 2/3$ . Jetzt müssen schon *zwei Drittel der Bevölkerung* *resistent* (bzw. später ggf. geimpft) geworden sein, bis daß lokale Ausbrüche von allein abklingen.

### Quiz

(1) Die Basis-Reproduktionszahl ist gleich \_\_\_\_\_. Sie ist interessant gerade \_\_\_\_\_. (2) Dann ist sie wichtig, weil der Vergleich der Basis-Reproduktionszahl mit \_\_\_\_\_ darüber entscheidet, ob \_\_\_\_\_. (3) Eine weitere Interpretation der Basis-Reproduktionszahl, die in der Presse häufig zu finden ist, ist die \_\_\_\_\_. (4) Für die wöchentliche Veränderung der Infektiösen schreiben wir  $\dot{I} = (\beta S/N - \gamma)I$ . Hier sieht man sofort, daß ein von selbst erlöschender Ausbruch dann vorliegt,

wenn \_\_\_\_\_ . (5) Es ist  $\dot{I} = (\beta S/N - \gamma)I$  eine Differentialgleichung: hier finden wir sowohl Ableitung  $\dot{I} = dI/dt$  als auch Niveau  $I$ . Am Anfang der Epidemie vereinfacht sich  $\dot{I}$  zu:  $\dot{I} = (\beta - \gamma)I$ . Diese ist eine einfachere Differentialgleichung. In ihr ist die wöchentliche Veränderung der Infektiösen dem Niveau der Infektiösen \_\_\_\_\_ . (7) Die Lösung dieser einfachen Differentialgleichung ist  $I(t) = e^{(\beta - \gamma)t}$ , eine \_\_\_\_\_ . Skizzieren Sie deren Graphen am Rand dieser Seite. Welche Eigenschaft des Graphen erinnert uns an die ersten Zahlen im Februar oder März 2020 zu Covid-19? (Hinweis: Deshalb gilt der corona-Virus als so außerordentlich gefährlich.) (8) Wenn wir die eben besprochene Exponentialfunktion logarithmieren, dann erhalten wir  $\ln I(t) =$  \_\_\_\_\_ . Der Logarithmus der Infektiösen ist also \_\_\_\_\_ in  $t$  . (9) Viele Graphiken, die anfangs in den Medien zu finden waren, trugen die logarithmierten Covid-19-Getesten gegen \_\_\_\_\_ ab. Die Idee war: War der gezeichnete Graph linear, dann ahnte man, daß \_\_\_\_\_ . (10) Epidemie-Politik scheint in zwei Varianten (oder eine Mischung dieser) daherkommen: Die eine Variante versucht \_\_\_\_\_ , die andere \_\_\_\_\_ .

## Übungen

7.1 (**Wichtige Kennziffern des SIR-Modells**) Füllen Sie die einfache Tab. (7.2) nach Möglichkeit aus. Interpretieren Sie alle Parameter bzw. Kennziffern. Erläutern Sie anschließend, wie sich Veränderungen in  $\beta$  bzw.  $\gamma$  auf den Verlauf der Epidemie auswirken.

$\beta$	$\gamma$	$r_0$	$\rho/N$	$1 - 1/r_0$
0.20	0.10			
0.15	0.10			
0.10	0.20			
0.30	0.10			

Tabelle 7.2: Basis-Reproduktionszahl, Herdenimmunität, Krise

7.2 (**Theorie vs. Empirie**) Das Modell ist nicht perfekt mit den empirisch verfügbaren Daten kongruen. Wir analysieren z.B. hier nur  $I(S)$ , während die Medien  $I(t)$  oder  $R(t)$  – oder auch  $D(t)$ , die Anzahl der am Virus Verstorbenen – zu publizieren meinen, mit  $t$  als einem Index für die aktuelle Woche. Im einzelnen:

- (a) Warum ist die Anzahl der in der laufenden Woche Covid-19-Positiv Getesteten nicht identisch mit den Neuinfektionen,  $\beta S(t)I(t)/N$ ?

- (b) Warum ist die Anzahl der in der laufenden Woche Covid-19-Positiv Getesteten nicht identisch mit den Infektiösen,  $I(t)$ ?
- (c) Warum ist die Gesamtzahl der bisher Covid-19 Positiv Getesteten nicht identisch mit den Infektiösen,  $I(t)$ ?
- (d) Warum ist die Anzahl der lt. Statistik am Virus Verstorbenen nicht identisch mit  $D(t)$ ?
- (e) Welche für die Ausbreitung wichtigen Eigenschaften des corona-Virus haben wir im hier vorgestellten SIR-Modell vernachlässigt?

7.3 (**Die Epidemiekurve**) Betrachten Sie unsere Epidemiekurve, Gl. (7.14). In einem Diagramm mit  $S$  auf der Abszisse und  $I$  auf der Ordinate wollen wir jetzt den Graphen der Funktion  $I(S)$  skizzieren – auch um zu begründen, wie er sich von Abb. (7.7) nach Abb. (7.8) verändert.

- (a) Skizzieren Sie zuerst den einfachen Graphen zu  $N - S$ , als erstem Summanden in Gl. (7.14).
- (b) Skizzieren Sie anschließend den Graphen zu  $\rho \ln(S/N)$ , als zweitem Summanden in Gl. (7.14).
- (c) Was gilt für den Punkt  $(S, I) = (N, 0)$ ?
- (d) Welche Steigungen haben beide Graphen wenn  $S = N$ ?
- (e) Wie verhält sich  $\rho \ln(S/N)$ , wenn  $S$  immer kleiner wird?
- (f) Der Graph zu  $I(S)$  ist lt. Gl. (7.14) die Summe der Graphen aus den ersten beiden Teilaufgaben. „Vermuten“ Sie jetzt den Graphen zu  $I(S)$ , indem Sie diese beiden Graphen „vertikal addieren“. Zeichnen Sie eine Skizze.

7.4 (**Krise**) Bestimmen Sie die Anzahl der Infektiösen  $I$  für  $S = \rho$ , d.h. die größtmögliche Anzahl Infektiöser.

7.5 (**Epidemie-Politik**) Die verschiedenen politischen Maßnahmen führen zu einem Rückgang von  $\beta$ , und damit zu einem Anstieg von  $\rho = \gamma N / \beta$ .

- (a) Überlegen Sie in Ihrem Diagramm aus der vorigen Aufgabe zuerst, wie ein Anstieg von  $\rho$  den Graphen von  $\rho \ln(S/N)$  verändert.
- (b) Überlegen Sie dann, wie sich folglich der Graph von  $I(S)$  anpaßt. (Hinweis: Hier zeigen Sie, daß ein Anstieg von  $\rho$  unsere Epidemiekurve tatsächlich nach „rechts“ und nach „unten“ verschiebt.)
- (c) Welche wichtige Implikation hat diese politikinduzierte Verschiebung der Epidemiekurve?

- 7.6 (**Epidemiekurve**) Betrachten Sie unsere Epidemiekurve, Gl. (7.14). Erläutern Sie, welchen Effekt eine Verdopplung von sowohl  $S$  als auch  $N$  auf die Anzahl der Infektiösen hat.



## 8 Evidenzbasierte Politik

### Be. 132: Vakzine und Impfung

Im Dezember 2020 erfolgte die erste Zulassung eines Impfstoffes gegen Covid-19, weitere Impfstoffe folgten seitdem. Die Möglichkeit des Impfens verlangt nach Anpassungen des SIR-Modells; Immunität kann jetzt auch ohne Infektion, durch Impfung, erworben werden. Eine formale Analyse fügen wir hier nicht mehr hinzu; aber intuitiv ist klar, daß Impfungen die Zahl der *susceptibles* senken. Senken sie sie hinreichend stark, wird Herdenimmunität erreicht, auch ohne daß ein großer Teil der Bevölkerung sich infizieren muß. Die Gesellschaft verkraftet eine „gewisse“ Anzahl von Impfgegnern. Es reicht – in der oben besprochenen Weise – schon, nur einen Anteil  $1 - 1/r_0$  zu impfen. Der verbleibende Anteil  $1/r_0$  darf Impfgegner sein. Die Gesellschaft muß keine (für Politiker im Wettbewerb der Demokratie nachteilige) Impfpflicht ausrufen. Allerdings: Ist der Anteil Impfgegner größer als  $1/r_0$ , geht diese Strategie nicht auf. Im November 2021 schätzte man, daß  $r_0 \approx 5$ . Damit war Herdenimmunität erst bei einem Anteil von 0.8 der Gesellschaft erreicht. Portugal erreichte damals allein schon mit seiner Impfquote von über 80% diesen Anteil; Deutschland mit seiner Impfquote von damals ca. 67% vermutlich (d.h. auch unter Hinzuziehen der bereits zwar Infizierten, aber eben nicht Geimpften) nicht.

### 8.1 Wirksamer Impfstoff?

#### Be. 133: Impfdurchbrüche: Antwort auf ... irrelevante Frage

Ein beliebtes Argument gegen die Impfung auf seiten der „Impfgegner“ ist der Hinweis auf den sog. „Impfdurchbruch“: Auch zweifach-Geimpfte würden Covid-19 bekommen und insbesondere auch hospitalisiert werden müssen. Auf den Stationen seien schon mindestens ein Drittel der Patienten Geimpfte. Nur: Im Kern trägt dieser Hinweis nichts zur Frage bei, „was die Impfung bringt“. Für die Frage, „was die Impfung bringt“, geht es nicht um den Anteil derer unter den Hospitalisierten, die geimpft sind (analog sagt man auch: ... nicht um die (bedingte) Wahrscheinlichkeit, als Hospitalisierter ( $H$ ) geimpft ( $I$ ) zu sein, also  $P(I|H)$ ). Es geht vielmehr bei der Frage nach dem Sinn der Impfung, i.e. um den Anteil der Hospitalisierten unter den Geimpften (analog: „um die (bedingte) Wahrscheinlichkeit, als Geimpfter hospitalisiert zu werden“, also  $P(H|I)$ ). *Diese*

Wahrscheinlichkeit legt den Grundstein für die für die Entscheidung relevante Frage nach dem „absoluten Ertrag“ oder Vorteil der Impfung:

$$\text{Ist } P(H|I) - P(H|\bar{I}) < 0?$$

oder alternativ auch in Gestalt einer „Änderungsrate dieses Ertrags“ oder des Vorteils der Impfung:

$$\text{Ist } \frac{P(H|I) - P(H|\bar{I})}{P(H|\bar{I})} < 0? \quad (8.1)$$

Daß dagegen  $P(I|H)$  in der öffentlichen Diskussion eine so prominente Stellung einnimmt, liegt daran, daß diese Wahrscheinlichkeit (dieser Anteil) leicht zu bestimmen ist – und natürlich an der häufigen Fehlinterpretation einfacher Kennziffern. Krankenhäuser zählen einfach ihre Hospitalisierten und erfragen unter ihnen den Anteil derer, die geimpft sind. Dieser Anteil ist typischerweise ziemlich hoch: während der Epidemie z.B. 30%, oder  $P(I|H) = 0.3$ . Anschließend wird diese Zahl von Chefärzten in *talk shows* u.ä. referiert.

#### Be. 134: Baum der Wahrscheinlichkeiten

Besprechen wir die diversen Ereignisse, die jedem Individuum im Zuge der Epidemie zwangsläufig bevorstehen. Es (i) kann geimpft sein (Ereignis Impfung,  $I$ ) oder nicht ( $\bar{I}$ ). Und (ii) es kann sich mit anschließend notwendig werdender Krankenhausbehandlung anstecken (Ereignis Hospitalisierung,  $H$ ) oder nicht (Ereignis  $\bar{H}$ ). Die zugehörigen Wahrscheinlichkeiten bezeichnen wir mit  $P(I) = 1 - P(\bar{I})$  bzw.  $P(H) = 1 - P(\bar{H})$ . Interessant ist jetzt die Interaktion zwischen Impfung und Hospitalisierung. Abb. (8.1) zeigt die möglichen Kombinationen in Gestalt eines *Baums*. Dort finden Sie entlang der ersten beiden Äste  $P(I) \approx 0.68$  und  $P(\bar{I}) \approx 0.32$  bereits eingetragen. Anschließend sehen Sie vier weitere Äste in Abhängigkeit vom Impfstatus. Die oberen beiden Äste repräsentieren die bedingten Wahrscheinlichkeiten, als Ungeimpfter hospitalisiert zu werden ( $P(H|\bar{I})$ ) oder eben nicht ( $P(\bar{H}|\bar{I})$ ). Die nachfolgenden unteren beiden Äste repräsentieren die bedingten Wahrscheinlichkeiten, sich als Geimpfter ins Krankenhaus begeben zu müssen ( $P(H|I)$ ) oder nicht ( $P(\bar{H}|I)$ ). – Ganz am Schluß, rechts im Diagramm finden Sie von oben nach unten die (als Produkt der zulaufenden Äste des jeweiligen Pfads) resultierenden Wahrscheinlichkeiten (Bevölkerungsanteile), ungeimpft und frei zu sein (rot mal grün, oder  $P(\bar{I}) \cdot P(\bar{H}|\bar{I}) = P(\bar{I} \cap \bar{H})$ ), ungeimpft und hospitalisiert zu sein (rot mal gelb, oder  $P(\bar{I} \cap H)$ ), geimpft und hospitalisiert zu sein (blau mal lila, oder  $P(I \cap H)$ ) und schließlich geimpft und frei zu sein (blau mal blau, bzw.  $P(I \cap \bar{H})$ ).

#### Be. 135: Die relevante bedingte Wahrscheinlichkeit

Heben wir noch einmal hervor, welche Wahrscheinlichkeit die eigentlich relevante, die interessante ist. Das ist wie besprochen  $P(H|I)$  (Be. (133)). Nach der

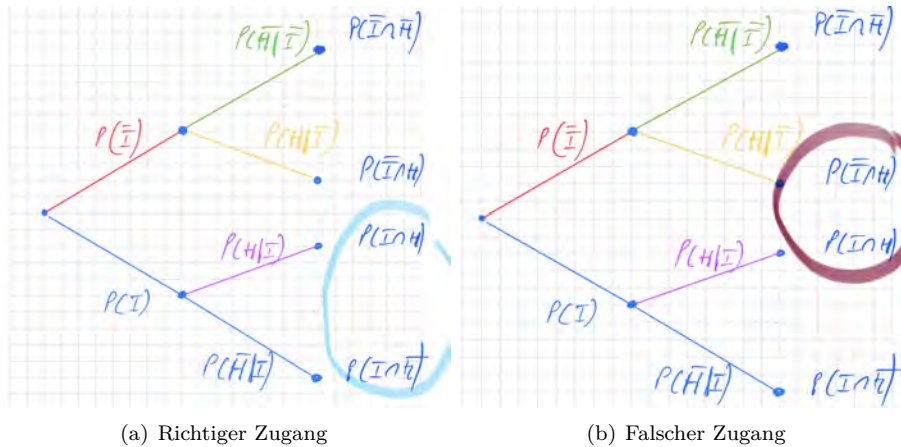


Abbildung 8.1: Baumdiagramm zu W.-Verteilung von Impfung und Hospital.

Definition der bedingten Wahrscheinlichkeit (oder alternativ auch nach unserer Interpretation als geeigneter Anteil) bestimmen wir sie als

$$P(H|I) = \frac{P(I \cap H)}{P(I)} = \frac{P(I \cap H)}{P(I \cap H) + P(I \cap \bar{H})}. \quad (8.2)$$

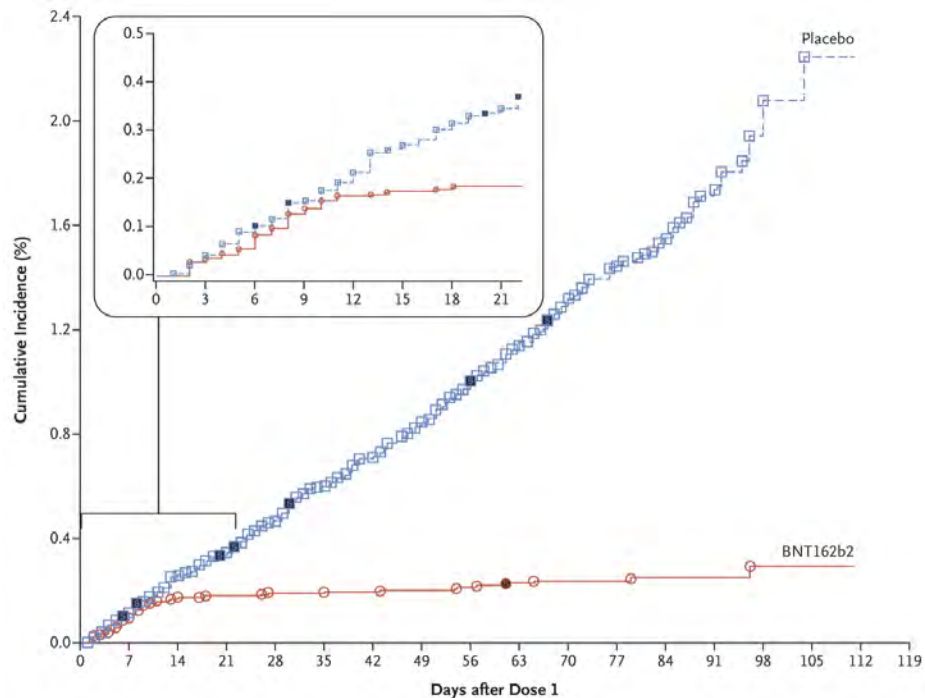
Die erste Gleichung oben ist also die Definition der bed. Wahrscheinlichkeit, als Geimpfter ins Krankenhaus zu müssen. Die zweite Gleichung nutzt zusätzlich im Nenner einfach aus, daß die Summe der Wahrscheinlichkeiten, als Geimpfter hospitalisiert zu werden oder eben nicht, gleich der Wahrscheinlichkeit ist, geimpft zu sein.<sup>1</sup> Oder denken Sie einfach einmal in Bevölkerungsanteilen: Die gesuchte Wahrscheinlichkeit ist doch einfach der Anteil der geimpft Hospitalisierten an allen Geimpften! Um  $P(H|I)$  zu bestimmen müssen, dürfen wir Gl. (8.2) zufolge (ausschließlich) die beiden Wahrscheinlichkeiten im hellblauen Kreis der Abb. (8.1a) nutzen.

### Be. 136: Die *irrelevante* bedingte Wahrscheinlichkeit

Die für den einzelnen dagegen insbesondere für seine Entscheidung, sich impfen zu lassen oder nicht, irrelevante Wahrscheinlichkeit ist die, als Hospitalisierter geimpft zu sein.<sup>2</sup> Diese Wahrscheinlichkeit ist mit analogen Überlegungen wie

<sup>1</sup>Hier nutzen wir die einzige Aussage aus der Statistik. Es gilt  $P(I) = P(I \cap H) + P(I \cap \bar{H})$ . Intuitiv ist das klar: Der Anteil der Geimpften in der Bevölkerung ist gleich dem Anteil der geimpft Hospitalisierten und der geimpft Nicht-Hospitalisierten.

<sup>2</sup>Vielleicht stimmt das nicht so ganz, wenn die Krankheitsverläufe im Krankenhaus angekommener geimpfter Patienten immer noch milder sind also die der dort ungeimpften Patienten. Dann wissen Ärzte immerhin, daß sie von den Geimpften mehr Resilienz erwarten dürfen.



Efficacy End-Point Subgroup	BNT162b2, 30 µg (N=21,669)		Placebo (N=21,686)		VE (95% CI) percent
	No. of participants	Surveillance time person-yr (no. at risk)	No. of participants	Surveillance time person-yr (no. at risk)	
Covid-19 occurrence					
After dose 1	50	4.015 (21,314)	275	3.982 (21,258)	82.0 (75.6–86.9)
After dose 1 to before dose 2	39		82		52.4 (29.5–68.4)
Dose 2 to 7 days after dose 2	2		21		90.5 (61.0–98.9)
≥7 Days after dose 2	9		172		94.8 (89.8–97.6)

Abbildung 8.2: Impfstoffwirksamkeit von BioNTech/Pfizer (Polack et al. (2020))

oben

$$P(I|H) = \frac{P(H \cap I)}{P(H)} = \frac{P(H \cap I)}{P(H \cap I) + P(H \cap \bar{I})}. \quad (8.3)$$

In Abb. (8.1b) konstruieren Sie diese Wahrscheinlichkeit (ausschließlich) mit den dort rot eingekreisten zwei Wahrscheinlichkeiten. Etwas anders: Hier nutzen Sie zwar den richtigen Zähler, aber den falschen Nenner! – Eine (eigentlich unerhebliche) Einschränkung hat diese Diskussion: die im Extrem notwendig werdende *triage*. Die *triage* ist die in den napoleonischen Schlachten eingeführte Dreiteilung („tri“) der Verwundeten in drei Gruppen, die der Aussicht ihrer erfolgreichen ärztlichen Rettung entsprechen. Manche fordern unterschiedliche Behandlungen für geimpft vs. ungeimpft Hospitalisierte – und Ärztevertreter führen mindestens ebenso häufig ganz andere Kriterien an. Aber auch hier ist

die Bestimmung von  $P(\bar{I}|H) = 1 - P(I|H)$  nicht informativ. Diese Wahrscheinlichkeit (bzw. dieser Anteil) dient der öffentlichen Empörung, vielleicht auch Abschreckung – fällt aber ohnehin mit zunehmender Impfquote  $P(I)$  immer weiter, bis daß am Schluß notwendigerweise nur noch (aber dann absolut überhaupt weniger) Geimpfte auf den Stationen liegen.

### Be. 137: Satz von Bayes

Jetzt haben wir fast schon den berühmten *Satz von Bayes* besprochen. Bayes setzt die beiden bedingten Wahrscheinlichkeiten (die relevante und die irrelevante) zueinander in Beziehung. Das gelingt auch uns leicht anhand der jeweils ersten Gleichung in Gl. (8.2) und Gl. (8.3): weil der Zähler auf den rechten Seiten beider (jeweils ersten) Gleichungen der gleiche ist. Er lautet in beiden Fällen:  $P(H \cap I) (= P(I \cap H))$ . Substituieren gibt:

$$P(H|I) = \frac{P(H)}{P(I)} P(I|H). \quad (8.4)$$

Der Satz von Bayes erlaubt das Transformieren der einen bedingten Wahrscheinlichkeit in die jeweils andere. Wir müssen dabei „nur“ den Korrekturfaktor  $P(H)/P(I)$  berücksichtigen. Kennen wir diesen Korrekturfaktor, können wir die irrelevante leicht in die relevante Wahrscheinlichkeit umrechnen. Ein Zahlenbeispiel mit einer entsprechenden Anwendung folgt weiter unten (Be. (141)).

### Be. 138: Diff-in-diff-Analyse

Die *difference-in-differences*-Analyse (auch diff-in-diff oder nur DD genannt) ist ein zentraler Baustein der Evaluation neuer Medikamente. Dort vergleicht man den Krankheitsverlauf derer, die das Medikament erhalten haben („treatment“-Gruppe), mit dem Krankheitsverlauf derer, die das Medikament nicht erhalten haben („control“-Gruppe). Im Vergleich der beiden Entwicklungen läßt sich das Medikament intuitiv besser beurteilen als nur durch den Blick auf den Verlauf der „treated“. Eine implizite Annahme hinter diesem Schluß ist, daß der Verlauf der „treated“ genau den Verlauf der „control“ genommen hätte, wären die „treated“ nicht behandelt worden. Diese Annahme heißt *common-trends*-Annahme. Oft unterscheiden sich die „treatment“- und die „control“-Gruppe auch in weiteren Eigenschaften als nur der, „treated“ zu sein oder nicht. Etwa: Ist das bessere Abschneiden der „treatment“-Gruppe vielleicht nur diese anderen Eigenschaften geschuldet? Aus diesem Grund wählen medizinische Evaluationsstudien das Design des sog. *randomized controlled trial* (RCT). Dabei sind die beiden Gruppen zwar nicht in jederlei Hinsicht identisch. Aber wegen der zufälligen Zuordnung der Teilnehmer in beide Gruppen sind ihre Unterschiede nicht mehr „systematisch“. Die große Evaluationsstudie des Impfstoffs von Bio-Tech/Pfizer<sup>3</sup> Insgesamt 43,448 Teilnehmer erhielten eine Injektion. Von diesen

<sup>3</sup>Vgl. Polack et al (2020), der dieser Strategie folgt.

bekamen 21,720 den Impfstoff, alle anderen eine Kochsalzlösung.<sup>4</sup> Anschließend verfolgte die Studie die Teilnehmer über ca. zwei Monate hinweg.

### Be. 139: Impfstoffwirksamkeit: Definition

Der in der Evaluationsstudie dokumentierte Erfolg von BioNTech/Pfizer schlägt sich auch in der Kennziffer der *Impfstoffwirksamkeit*  $E$  (auch *efficacy* oder *relative risk reduction*) nieder. Die ist definiert als:

$$E = \frac{\text{Covid-19 in Kontrollgruppe (je Tn.)} - \text{Covid-19 in Treatmentgruppe (je Tn.)}}{\text{Covid-19 in Kontrollgruppe (je Tn.)}}$$

Sie mißt innerhalb der Studiengruppe, um wieviel Prozent die Wahrscheinlichkeit, zu erkranken, bei Geimpften geringer ausfällt als bei Nicht-Geimpften. Noch kürzer: Um wieviel ist die Wahrscheinlichkeit, zu erkranken, durch die Impfung geschrumpft? Diese Frage ist für uns jetzt aber eigentlich nicht überraschend: Wir hatten sie uns doch schon oben in Ungleichung (8.1) als Kriterium für die Attraktivität der Impfung überlegt. (Sie müssen nur „Ansteckung“ Covid-19 mit „Krankenhausaufenthalt“  $H$  vertauschen und ein Minuszeichen vor den Ausdruck auf der linken Seite der Ungleichung schmuggeln, damit er grundsätzlich positiv wird und sich positiv besetzt als „Wirksamkeit“ interpretieren läßt.)

### Be. 140: Impfstoffwirksamkeit: Pfizer/BionTech

Abb. (8.2) zeigt sehr eindrucksvoll, wie die Inzidenzen bei Geimpften und Ungeimpften (Placebo- oder Kontrollgruppe) ab 7 Tagen nach Verabreichen der zweiten Dosis auseinanderzudriften zu beginnen. Nach 105 Tagen lag die kumulative Inzidenz der 21699 Teilnehmer der Treatmentgruppe bei 9, die der 21686 Teilnehmer der Kontrollgruppe bei 172. Daher ist die Impfstoffwirksamkeit

$$E = \frac{172/21686 - 9/21669}{172/21686} \approx 0.95. \quad (8.5)$$

Der Anteil Covid-19-Erkrankter war damit in der Treatmentgruppe um fast 95% kleiner als in der Kontrollgruppe! – Die hier beschriebene Methodik, die eigentlich zur Beurteilung neuer Medikamente – und aktuell gerade neuer Impfstoffe – gemeint ist, hält seit einigen Jahren interessanterweise auch Einzug in die Evaluation wirtschaftspolitischer Maßnahmen, und das ist ein weiterer Grund, warum wir sie hier besprechen. Dort ist ein „treatment“ ein wirtschaftspolitisches (Förder-, Steuer o.ä.) Instrument, dem die Bewohner einer Region ausgesetzt sind, die Bewohner anderer Regionen dagegen nicht: hier liegt der Bezug zur Regionalökonomie. Regionen sind „Labor“ der Wirtschaftswissenschaftlerin.

<sup>4</sup>Um subjektive Beurteilungen bzw. sog. „placebo“-Effekte auszuschließen, war diese Zufallsauswahl den Teilnehmern nicht bekannt.

**Be. 141: Impfstoffwirksamkeit: Deutschland**

Bestimmen wir zuletzt die Impfstoffwirksamkeit jetzt auch noch aus den uns verfügbaren Daten der vielen Millionen Bewohner Deutschland insgesamt (statt nur zu den relativ wenigen Teilnehmern der BioNTech/Pfizer-Studie) – auch wenn der Begriff Impfstoffwirksamkeit dann nicht mehr ganz exakt ist. Dazu überführen wir die uns bekannte bedingte Wahrscheinlichkeit  $P(I|H)$  in die uns interessierende Wahrscheinlichkeit  $P(H|I)$  nach dem Satz von Bayes in Gl. (8.4). Laut Robert-Koch-Institut ist die Zahl der Hospitalisierten am 26. November 2021 gerade 5.97 je 1,000 Ew. also  $P(H) \approx 0.00006$ . Und von  $P(I)$  wissen wir, daß sie zu diesem Zeitpunkt bei ca. 0.68 liegt. Damit folgt:

$$P(H|I) = \frac{P(H)}{P(I)} P(I|H) = \frac{0.00006}{0.68} \cdot 0.3 \approx 0.0026 \%. \quad (8.6)$$

Gleichzeitig ist die Wahrscheinlichkeit, als Ungeimpfter hospitalisiert zu werden, mit analogen Überlegungen gleich

$$P(H|\bar{I}) = \frac{P(H)}{P(\bar{I})} P(\bar{I}|H) = \frac{0.00006}{0.32} \cdot 0.7 \approx 0.0131 \%. \quad (8.7)$$

Damit ist die „Impfstoffwirksamkeit“ in unserer groben Rechnung

$$E = \frac{P(H|\bar{I}) - P(H|I)}{P(H|\bar{I})} = \frac{0.0131 - 0.0026}{0.0131} \approx 0.802 \quad (8.8)$$

Der Impfstoff reduziert die Wahrscheinlichkeit, zu erkranken, um 80%. Das sind nicht die 95% der BioNTech-Studie, aber erstens waren unsere Überlegungen ja nur sehr grob, und zweitens bezog sich die Zulassungsstudie nur auf das Zwei-Monats-Fenster nach zweiter Dosis. Unsere Überlegungen gelten für den Durchschnitt der Bevölkerung und für deutlich längere Zeiträume. Für einzelne Bevölkerungsgruppen – z.B. für Jüngere vs. Ältere – wird diese Zahl anders ausfallen.

## 8.2 Evidenzbasierte Evaluation

**Be. 142: Regionale Lohnentwicklung**

Die bisherige Diskussion bezog sich den Unterschied der Covid-19-Inzidenz zwischen Treatment- und Kontrollgruppe. Aus einer regionalen Perspektive unterscheiden wir jetzt zwischen zwei Regionen, deren eine eine Treatment (eine regionalpolitische Maßnahme, einen natürlichen Schock wie ein Extremwetterereignis oder ein Erdbeben) erfährt, deren andere dagegen nicht. Dabei interessieren uns der Effekt eines Treatments auf den Lohn. Wir beobachten dazu den Lohn in zwei Regionen namens  $i = A, B$  zu jeweils zwei Zeitpunkten namens  $t = 0, 1$ .

Wir haben damit insgesamt vier Beobachtungen zu Löhnen. Jetzt sei ein Schock eingetreten: z.B. ein Erdbeben. Dieser Schock hat die beiden folgenden Eigenschaften: Er ereignet sich zwischen Zeitpunkt 0 und Zeitpunkt 1. Und er trifft Region A, aber nicht Region B. Damit wird Region A zur *Treatment*-Region, Region B zur *Control*-Region. Gleichzeitig wird Zeitpunkt 0 zum *prä-Schock*-Zeitpunkt, Zeitpunkt 1 dagegen zum *Post-Schock*-Zeitpunkt. Die beiden Regionen seien übrigens Vergleichs-, aber nicht über Migration verknüpfte Regionen. Was in der einen Region dient als „benchmark“ für das, was sich in der anderen ereignet; sie liegen nicht so dicht beieinander, daß Migration im Anschluß an den Schock zwischen ihnen mittelbar zu einer „Verteilung“ der Wirkungen des Schocks auf beide Regionen führt. (Dann wäre Region B keine optimale Kontrollregion.)

**Be. 143: Einfache Vergleiche – und ihre Defizite**

Wir könnten die Auswirkungen eines Schocks (Treatments) durch zweierlei Arten von Vergleichen schätzen. Wir könnten den Lohnunterschied zwischen beiden Regionen nach dem Schock (Treatment) heranziehen:

$$y_{A2} - y_{B2}. \quad (8.9)$$

Oder: Wir könnten die Lohnveränderung nur in der vom Schock (Treatment) betroffenen Region heranziehen:

$$y_{A2} - y_{A1} \quad (8.10)$$

Beides sind häufig benutzte Konzepte. Beide Konzepte haben allerdings ein jeweils eigenes Problem. Der interregionale Lohnunterschied post-schock (8.9) ist nicht mehr so beeindruckend, wenn es einen interregionalen Lohnunterschied bereits prä-schock gegeben hat. Sicher wird man nur einen *Zuwachs des Lohnunterschieds* als Ausdruck des Schocks gelten lassen. Und die in der vom Schock betroffenen Region A beobachtete Lohnveränderung (8.10) ist nicht mehr so überzeugend, wenn es auch eine Lohnänderung in der vom Schock gar nicht betroffenen Region B gegeben hat. Sicher wird man nur das *Mehr an Lohnänderung* in der betroffenen Region als „kausalen Ausdruck“ des Schocks gelten lassen. – Beide Einwände gehen auf den gleichen fundamentalen Einwand zurück: Regionale Unterschiede sind nie nur das Ergebnis der Variation einer einzelnen exogenen Variable. Sie reflektieren die *simultane Variation mehrerer, sogar vieler exogener Variable*.

**Be. 144: Unterschied in Änderungen = Änderung der Unterschiede**

Diese Überlegungen führen uns auf die folgenden neuen „korrigierten“ Konzepte: Wir ziehen nicht einfach den Lohnunterschied nach dem Schock heran, sondern nur den *Zuwachs* des Lohnunterschieds. Also:

$$\left( y_{A2} - y_{B2} \right) - \left( y_{A1} - y_{B1} \right) \quad (8.11)$$



Oder: Wir ziehen nicht einfach die Lohnveränderung in Region *A* heran, sondern nur die um die Lohnveränderung in Region *B* korrigierte Lohnveränderung der Region *A*. Also:

$$(y_{A2} - y_{A1}) - (y_{B2} - y_{B1}) \quad (8.12)$$

Beide Konzepte sind überzeugender als die einfachen Differenzen aus der vorigen Bemerkung. Beide Konzepte sind darüber hinaus ... tatsächlich *identisch*. Das sehen Sie sofort selbst, wenn Sie die „Veränderung des Unterschieds“ aus Gl. (8.11) mit „Unterschied der Veränderungen“ aus Gl. (8.12) miteinander vergleichen.<sup>5</sup> Jetzt wird mit diesem Vokabular auch klar, warum die Literatur hier von *difference-in-differences* spricht. Für praktische Anwendungen interessant ist, wie man (8.12) möglichst schnell ausrechnet. Schwierig ist das auf direktem Weg natürlich nicht. Trotzdem machen wir jetzt einen Umweg (von dem wir erst später sehen, daß er sich lohnt). Wir zeigen, daß der Ausdruck (8.12) sich als Spezialfall des OLS-Schätzers einstellt.

#### Be. 145: Diff-in-Diff: „intuitive“ Anwendungen in diesem Skript

Hier sind Beispiele dieses Skripts, in denen wir „intuitiv“ bereits den Unterschied in der zeitlichen Entwicklung von zwei (oder sogar mehr) Regionen herangezogen haben: Wir besichtigen nicht naiv ...

- (Peso-Krise): ... die Lohnentwicklung in Kalifornien. Wir besichtigen den Unterschied zwischen der Lohnentwicklung Kaliforniens und der New Yorks (Be. (47) und Abb. (2.8)).
- (Große Depression): ... die Entwicklung der Arbeitslosenquote in München. Wir besichtigen den Unterschied zwischen der Arbeitslosenentwicklung Münchens und der anderer (süd-)deutscher Kreise (Abb. (4.1)).
- (Covid-19): ... die Immobilienmarktindikatoren im Zentrum. Wir besichtigen den Unterschied zwischen der Entwicklung dieser Indikatoren und der in Vorstädten oder sogar Vorvor-Städten.

Und hier sind Beispiele dieses Skripts, in denen wir „intuitiv“ die zeitliche Entwicklung eines Unterschieds zwischen zwei (oder sogar mehr) Regionen herangezogen haben: Wir besichtigen nicht naiv ...

<sup>5</sup>Vielleicht sind Sie mit den Methoden der Differentialrechnung so gut vertraut, daß Ihnen folgendes alternatives Argument einleuchtet – und damit Ihre Intuition stützt. Es sei  $t$  der Zeitindex und  $i$  der Regionsindex. Dann ist die Veränderung des Lohns in der Zeit die 1. partielle Ableitung  $\partial y / \partial t$ . Und die Veränderung dieser Veränderung hinsichtlich der Region ist die 2. gemischte Ableitung  $\partial^2 y / \partial t \partial i$ . Bei der zweiten gemischten Ableitung spielt es bekanntlich keine Rolle, in welcher Reihenfolge Sie ableiten. Es ist einfach  $\partial^2 y / \partial t \partial i = \partial^2 y / \partial i \partial t$ . (Das ist Youngs Theorem.) Und diese zweite gemischte Ableitung ist gerade der „difference-in-differences“.

- (Lohn-Verteilung der USA): ... die Lohnunterschiede amerikanischer Regionen im Jahr 2009, sondern den Unterschied zwischen der Lohnverteilung 2009 und der von 1964 (Abb. (2.9))
- (Bevölkerungs-Verteilung der USA): ... die Bevölkerungsverteilung der USA im Jahr 1992, sondern den Unterschied der Bevölkerungsverteilung 1992 zu der zu 1976 (Abb. (6.1)).

Ein Nachtrag zur „Kontrollgruppe“, also zur Region, die kein Treatment erfährt: Sie könnten meinen, das eigentliche Geschehen spiele in der Treatment-Gruppe. Aber das stimmt ja eigentlich nicht. In der Entwicklung der Treatment-Gruppe spiegelt sich *alles*, was sich im Zeitablauf ändert: Das Treatment, aber eben auch das Hintergrundrauschen der Konjunktur, des Wachstums, von Krisen und Kriegen, u.v.a.m. In der Entwicklung der Kontroll-Gruppe spiegelt sich dagegen (wenn es eine gute Kontrollgruppe ist) *alles außer dem Treatment*. Die Kontrollgruppe ist also ein kleines bißchen präziser. Und ohnehin ist es ja die Differenz zwischen beiden Gruppen, die überhaupt erst informativ ist, weil überhaupt erst sie den Treatment-Effekt freilegt.

#### Be. 146: Datensatz

Versammeln wir zuerst alle vier Löhne in einer Liste, in einem Vektor  $y$ ,

$$y = \begin{pmatrix} y_{A1} \\ y_{A2} \\ y_{B1} \\ y_{B2} \end{pmatrix}. \quad (8.13)$$

Gleichzeitig wollen wir uns merken, wo und wann diese Beobachtungen gemacht wurden. Hierfür notieren wir zwei regionale Indikator- oder Dummy-Variablen

$$d_A = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad d_B = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad (8.14)$$

sowie zusätzlich zwei weitere zeitpunktbezogene Indikator-Variablen

$$d_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad d_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}. \quad (8.15)$$

Interessant ist schließlich der Vektor  $d_A d_2$ . Es verweist auf diejenigen Beobachtungen, die tatsächlich nur vom Schock betroffen sind. Diese Beobachtungen

sind ... ja nur eine einzige Beobachtung: die in Stadt  $A$  zum Zeitpunkt 2. Jene identifizieren wir mit Hilfe der Indikatorvariable

$$d_A d_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \quad (8.16)$$

Regressieren wir jetzt  $y$  auf  $d_A, d_B, d_2$  sowie  $d_A d_2$ . Warum? Weil uns das eine Schätzung des Ausdrucks 8.12 liefert wird, wie wir in Kürze sehen werden.

**Be. 147: Frisch-Waugh-Lovell-Theorem (Residualregression)**

Wir wollen die Koeffizienten  $\alpha_1, \alpha_2, \beta$  sowie  $\gamma$  der folgenden Gleichung

$$y = \alpha_1 d_A + \alpha_2 d_B + \beta d_2 + \gamma d_2 d_A + \varepsilon \quad (8.17)$$

schätzen. „Schätzen“ heißt: wir suchen  $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}$  und  $\hat{\gamma}$  nach dem Kleinst-Quadrate-Kriterium. Das ist zugegeben erst einmal nicht so einfach: Wir haben vier Regressoren, nicht zwei oder sogar nur einen. Aber jetzt nutzen wir einen Trick bzw. eine berühmte Aussage der Ökonometrie, das Frisch/Waugh/Lovell-Theorem (Greene (1997, Kap. 6), Goldberger (1991, Kap. 17.3)). Ihm zufolge dürfen wir stufenweise vorgehen. Wir werden (i) zuerst  $d_2$  als auch  $d_2 d_A$  jeweils auf  $d_A$  und  $d_B$  regressieren, um (ii) sodann  $y$  auf die Residuen der beiden Regressionen aus (i) zu regressieren: Die in Schritt (ii) geschätzten zwei Koeffizienten sind dann gerade die gesuchten  $\hat{\beta}$  und  $\hat{\gamma}$ . Diese Schritte gehen wir in den folgenden Bemerkungen (148) und (149). Sie sind nicht wirklich schwierig, weil sie jeweils immer nur zwei Regressoren erfordern, und mit zwei Regressoren umzugehen ist okay. Anschließend halten Sie ein zentrales Resultat der modernen *evidenzbasierten Politikevaluation* in den Händen (Be. (150)).

**Be. 148: Hilfsregressionen**

Im ersten Schritt regressieren wir  $d_2$  auf  $(d_A, d_B)$ . Hierfür nutzen wir die bekannte Formel des OLS- bzw. KQ-Schätzers, i.e.  $(X'X)^{-1}X'y$ . In Üb. (8.4) zeigen Sie, daß Sie als Residuen dieser ersten Hilfsregression gerade

$$0.5 \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \quad (8.18)$$

erhalten. Anschließend regressieren wir  $d_2 d_A$  auf  $(d_A, d_B)$ . In Üb. (8.5) zeigen Sie, daß Sie als Residuen dieser zweiten Hilfsregression gerade

$$0.5 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad (8.19)$$

erhalten.

**Be. 149: Residualregression**

Im zweiten Schritt regressieren wir  $y$  auf die Residuen der beiden Hilfsregressionen (8.18) und (8.19) aus Be. (148). Das heißt konkret, daß die Regressor-Matrix

$$\tilde{X} = 0.5 \begin{pmatrix} -1 & -1 \\ 1 & 1 \\ -1 & 0 \\ 1 & 0 \end{pmatrix} \quad (8.20)$$

lautet, während der Regressand unverändert  $y$  aus Gl. (8.13) ist. Mit Hilfe von OLS bestimmen wir jetzt die Koeffizienten der *ursprünglichen* Regressionsgleichung (8.17) durch eine Regression von  $y$  auf  $\tilde{X}$ , also

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'y. \quad (8.21)$$

Der Anschaulichkeit halber machen wir das „zu Fuß“, und schrittweise, in den folgenden Bemerkungen. Vorher halten wir fest, daß die Formel (8.21) „machbar“ ist. Die Matrix  $\tilde{X}$  ist  $2 \times 2$ , der Vektor  $y$  ist  $2 \times 1$ . Nach den Regeln der Matrixmultiplikation hat  $(X'X)^{-1}X'y$  dann wiederum die Dimensionen  $2 \times 1$ .

**Be. 150: Mit Residualregression zum Difference-in-Differences**

Wir bestimmen  $(X'\tilde{X})^{-1}\tilde{X}'y$  in vier Schritten. Wir bestimmen zuerst  $\tilde{X}'\tilde{X}$ , dann die Inverse dieses Produkts, dann  $\tilde{X}'y$  und schließlich  $(X'\tilde{X})^{-1}\tilde{X}'y$ , das Produkt aus zweitem und drittem Schritt. – Es ist also

$$\tilde{X}'\tilde{X} = 0.5 \begin{pmatrix} -1 & 1 & -1 & 1 \\ -1 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 \\ 1 & 1 \\ -1 & 0 \\ 1 & 0 \end{pmatrix} 0.5 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}. \quad (8.22)$$

Mit der einfachen Regel der Matrix-Inversion für  $2 \times 2$ -Matrizen (Sydsaeter/Hammond (1995)) folgt

$$(\tilde{X}'\tilde{X})^{-1} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}. \quad (8.23)$$

Weiter ist

$$\begin{aligned} \tilde{X}'y &= 0.5 \begin{pmatrix} -1 & 1 & -1 & 1 \\ -1 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} y_{A1} \\ y_{A2} \\ y_{B1} \\ y_{B2} \end{pmatrix} \\ &= 0.5 \begin{pmatrix} (y_{A2} - y_{A1}) + (y_{B2} - y_{B1}) \\ (y_{A2} - y_{A1}) \end{pmatrix} \end{aligned} \quad (8.24)$$

Aber dann ist durch Multiplizieren von (8.23) mit (8.24) schließlich

$$(\tilde{X}'\tilde{X})(\tilde{X}'y) = \begin{pmatrix} \cdots \\ (y_{A2} - y_{A1}) - (y_{B2} - y_{B1}) \end{pmatrix}. \quad (8.25)$$

Jetzt sehen Sie leicht: Der Koeffizient in der zweiten Zeile auf der rechten Seite von Gl. (8.25) ist *gerade ... der gesuchte difference-in-difference aus (8.12)!* Fassen wir zusammen: Der geschätzte Koeffizient des Interaktionsterms  $d_2d_A$  in Gleichung (8.17) ist der *difference-in-difference*-Ausdruck (8.12). Kommen wir zum Ausgangspunkt zurück: Natürlich hätten wir den Difference-in-Difference (DiD) leichter bestimmen können. Aber daß der DiD auch in Gestalt des geschätzten Koeffizienten einer geeignet gewählten Regressionsgleichung auftritt, erlaubt uns Verallgemeinerungen, für die die einfache Formel (8.12) sich nicht mehr eignet. Etwa können wir den Einfluß *weiterer Größen* auf den Lohnsatz berücksichtigen. (Das ist besonders interessant, wenn *control*- und *treatment*-Gruppe sich nicht nur in dem *treatment*, sondern in weiteren (beobachtbaren) Aspekten unterscheiden. Dann nämlich könnte man letztere in das Modell aufnehmen und so hoffen, doch wieder den *treatment*-Effekt bestimmen zu können.)

## 8.3 Ausblick

### Be. 151: Impfung: Positiver Externer Effekt

Die Impfung ist ein positiver externer Effekt. In der Diskussion der Medien dominiert die Diskussion der individuellen Kosten-Nutzen-Analyse. Dort wird überlegt, ob sich die Impfung für den einzelnen lohnt. Aus gesellschaftlicher Perspektive relevant ist dagegen, ob sich die Impfung für die Gesellschaft lohnt. Der soziale Grenzertrag der Impfung ist größer als der individuelle: Die Impfung transformiert einen bisherigen *susceptible* in einen *removed* – und unterbricht so die mögliche Infektionskette an dieser Stelle. Wir kennen schon die Stärke dieses positiven externen Effekts. Dieser ist die effektive Reproduktionszahl, als Anzahl an in Abwesenheit der Impfung sonst infizierten Anfälligen (Be. (121)). Da die effektive Reproduktionszahl während der Epidemie typischerweise größer als Eins ist, schützt jede Impfung nicht nur den Geimpften, sondern sogar noch strikt mehr als einen weiteren *susceptible* – also strikt mehr als zwei Individuen. *Das* ist das zentrale Argument zugunsten einer Impfpflicht – oder zugunsten eines Impfgebots. Im Kontext positiver externer Effekte ist staatlicher Zwang daher (Koordination der Gesellschaftsmitglieder) eine mögliche sinnvolle Intervention. Eine alternative Lösung hierzu wäre eine (Pigou-)Subvention. Auch diese Lösung wurde versucht: „Bratenessen umsonst neben Impfbus in Mecklenburg-Vorpommern“; Impfprämie in New York; „Impfen ohne Impfausweis und Ausweisungspflicht in Berlin“ etc.

**Be. 152: Entscheidung gegen die Impfung**

Rekapitulieren wir drei populäre Einwände gegen eine Impfung:

- „Nicht jeder steckt sich zwangsläufig an“. Im SIR-Modell ist das richtig (Be. (130)). Im Zuge der Epidemie gelingt es einem kleinen Anteil – er ist kleiner als  $\rho/N = 1/r_0 \approx \frac{1}{5}$  – dem Virus zu entkommen. Aber ob gerade ich diejenige bin, der das gelingt? Und ist unser allereinfachste Modell das richtige in einem Kontext von Reinfizierbarkeit selbst nach Impfung und in der Präsenz immer neuer Virus-Mutationen (vgl. aktuell „Omikron“).
- „Die Wahrscheinlichkeit, daß ich ins Krankenhaus muß, ist klein.“ Für Jüngere ist diese Aussage richtig. Aber selbst für Jüngere ist diese Aussage wenig produktiv. Sie verstellt den Blick auf die eigentlich interessante Eigenschaft der Impfung (auch für Jüngere): Die hohe Impfstoffwirksamkeit von Pfizer/BioNTech signalisiert, daß die Impfung die (immerhin positive) Wahrscheinlichkeit einer Krankenhaus-Behandlung drastisch reduziert: um ca 80% im Durchschnitt der Bevölkerung (Be. (141)) oder sogar um ca. 95% im Durchschnitt der Studienteilnehmer von Pfizer/BioNTech (Be. (140)).
- „Auch Geimpfte enden im Krankenhaus.“ Und auch diese Aussage ist richtig. Aber eine wahre Aussage, die keine Antwort auf die eigentlich im Raum stehende Frage ist, ob sich die Impfung lohnt, ist sicher kein Einwand gegen die Impfung (Be. (136)). Vielleicht hilft folgendes Beispiel: Wenn Sie mich fragen würden, ob es morgen regnet: Würden Sie als Antwort die (wahre) Aussage „Wenn es regnet, wird es naß.“ akzeptieren?

Über mögliche Nebenwirkungen der Impfung (bzw. auch nur die Sorge hierüber) haben wir zugegeben in diesem und dem vergangenen Kapitel nichts gesagt. Mit etwas zeitlichem Abstand wäre auch noch einmal ein sachlicher Rückblick auf die dann vorliegende Evidenz zu Nebenwirkungen/keinen Nebenwirkungen interessant.