

# Seltene Desaster, Corona und verzerrte Stichproben

Prof. Dr. Joachim Grammig und Dr. Jantje Sönksen

Lehrstuhl für Statistik und Ökonometrie, empirische Wirtschaftsforschung

Tübingen, 14. Juli 2020

- ▶ Corona-Maßnahmen mit massiven wirtschaftlichen Konsequenzen.  
Wie gefährlich/tödlich ist das Virus?
- ▶ Pandemie als ein weltweites Desaster, ein *schwarzer Schwan*.  
Wird so etwas in ökonomischen Modellen berücksichtigt?  
Wenn nein, warum nicht? Was sind die Konsequenzen?
- ▶ Gemeinsames Phänomen: Selektionsverzerrte Daten.  
Wir erläutern das Problem und präsentieren eine Lösungsidee.

*“Erschwerend kommt hinzu, dass man die Schlussfolgerungen in der aktuellen Situation auf besonders unsicheres Datenmaterial gründen muss. So hat die Zahl der getesteten Infizierten nur bedingt etwas mit der Zahl der tatsächlichen Infizierten zu tun, weil Menschen mit wenigen oder gar keinen Symptomen bislang in den seltensten Fällen getestet werden”*

25. März 2020, “Statistische Konzepte und ihre Grenzen”

<https://www.rwi-essen.de/unstatistik/101>

*“Nur eine hinreichend großangelegte repräsentative Panelstichprobe von Personen, die sich regelmäßig in kurzer Frequenz einem Test unterziehen, kann das zentrale Problem der mangelnden Kenntnis der Dunkelziffer und damit der wahren Ansteckungsgefahr lösen”*

30. April 2020, “Die Reproduktionszahl und ihre Tücken”

<https://www.rwi-essen.de/unstatistik/102/>

# Aber es gibt doch gute Daten zu Corona\*?

\*Coronavirus SARS-CoV-2 und die ausgelöste Infektionskrankheit Covid-19

Robert-Koch-Institut

[https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/  
Fallzahlen.html](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html)

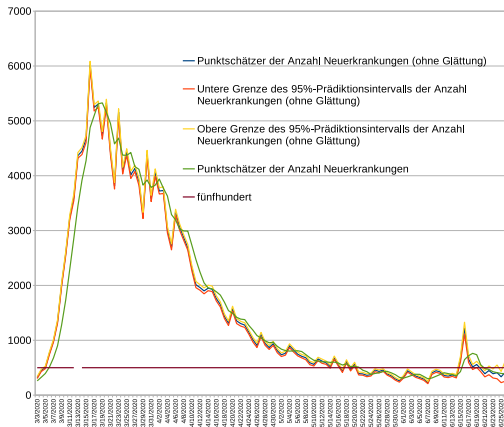
WHO

<https://covid19.who.int/info>

# Aber es gibt doch gute Daten zu Corona?

RKI Nowcasting: die besten Daten zur Verfolgung des Infektionsgeschehens in Deutschland

[https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/Projekte\\_RKI/Nowcasting\\_Zahlen.xlsx?\\_\\_blob=publicationFile](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Projekte_RKI/Nowcasting_Zahlen.xlsx?__blob=publicationFile)



# Aber es gibt doch gute Daten zu Corona?

Robert-Koch-Institut

[https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/Fallzahlen.html](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html)

WHO

<https://covid19.who.int/info>

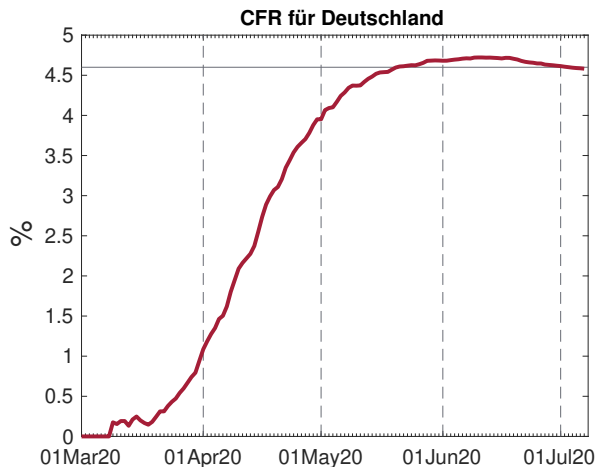
- ▶ Daten sind sehr gut, aber dennoch Selektionseffekte
- ▶ Weder Vollerhebung noch zufällig ausgewählte Covid-19-Infizierte
- ▶ Nur solche Infizierte kommen in die Stichprobe (werden getestet), die hinreichend starke Symptome haben
- ▶ Sinnvoll bei knappen Test- und Behandlungskapazitäten
- ▶ Nicht hilfreich zur Beantwortung der Frage: *Wie tödlich ist Covid-19?*

# Wie tödlich ist Covid-19? “Fake News” und “Facts”

- ▶ Wie wahrscheinlich ist es, dass ein Corona-Infizierter mit/an Covid-19 stirbt? Infection-Fatality-Rate: IFR
- ▶ “99% vollkommen harmlose Covid-19 Verläufe” (D. Trump)  
“Covid-19 wie eine leichte Grippe” (J. Bolsonaro)  
Grippesterblichkeit: IFR  $\approx 0,1\%$  (1 von 1000)
- ▶ RKI (13. Juli 2020)  
198,963 positiv auf SARS-Cov-2 getestete Personen  
9,064 Todesfälle mit/durch Covid-19  
(Symptomatic) Case-Fatality-Rate (CFR)  
 $\frac{8948}{192556} \times 100 = 4,56\%$  (46 von 1000)
- ▶ Covid-19-Sterblichkeit  $\approx 46$  mal höher als Grippe?



# Die CFR für Deutschland hat sich stabilisiert



Quelle: <https://covid19.who.int/info>

- ▶ CFR für andere Länder (<https://covid19.who.int/info> 8.7.20)
  - USA: 4,4%
  - China: 5,5%
  - Schweden: 7,4% (ca. 50% aus Seniorenheimen)
  - Italien: 14,4%
  - Vereinigtes Königreich: 15,5%
  - Island: 0,53%
  - Brasilien: 4,0%
- ▶ IFR sollte nicht stark variieren, es ist der gleiche Erreger!
- ▶ Temporäre Überlastung (Italien)
  - Bevölkerungsstruktur
  - Testpraxis
- ▶  $CFR \neq IFR$  (Infection-Fatality-Rate)
  - Nicht jeder Infizierte wird getestet: Hinreichend Symptome / eher schwere Verläufe

- ▶ Diamond Princess IFR  $\approx 1,8\%$   
alle getestet, 712 infiziert, 13 Covid-19-Tote  
Selektionseffekt: mittl. Alter 60 Jahre (Median)  
[https://de.wikipedia.org/wiki/Schiffe\\_unter\\_Quarant%C3%A4ne\\_in\\_der\\_COVID-19-Pandemie#Diamond\\_Princess](https://de.wikipedia.org/wiki/Schiffe_unter_Quarant%C3%A4ne_in_der_COVID-19-Pandemie#Diamond_Princess)
- ▶ Gangelt-Studie IFR  $\approx 0,37\%$   
Stichprobe, 919 getestet, 15% positiv, 7 Covid-19-Tote  
Altersstruktur repräsentativ? (Altenheime wurden verschont)  
<https://www.medrxiv.org/content/10.1101/2020.05.04.20090076v2>
- ▶ CFR Island  $\approx 0,53\%$   
10 Covid-19-Tote, 1.873 Infizierte, 364.134 Einwohner  
Selektionseffekt? Schutz von Risikogruppen?
- ▶ Kleine Fallzahlen und Selektionseffekte

# Vorschlag einer alternativen Schätzung IFR

- $M$  infizierte Person stirbt an/mit Covid-19
- $M-$  infizierte Person überlebt Covid-19
- $T$  Corona-Test durchgeführt und positiv
- $T-$  Corona-Test nicht durchgeführt
- $C$  Person ist mit SARS-CoV-2 infiziert

$W(E|A)$  Wahrscheinlichkeit von Ereignis  $E$ , gegeben Ereignis  $A$

$$\text{IRF} = W(M|C)$$

$\text{CRF} = W(M|C \text{ und } T)$  für DE mit RKI Daten gut abschätzbar  $\approx 4,5\%$

$$W(M|C) = W(M|C \text{ und } T) \times W(T|C) + W(M|C \text{ und } T-) \times W(T-|C)$$

(Satz der Totalen Wahrscheinlichkeit)

Annahme:  $W(M|C \text{ und } T-) \approx 0$  für DE OK, generell nicht ....

# Vorschlag einer alternativen Schätzung IFR

- $M$  infizierte Person stirbt an/mit Covid-19
- $T$  Corona-Test durchgeführt und positiv
- $C$  Person ist mit SARS-CoV-2 infiziert

$$\begin{aligned} IFR = W(M|C) &\approx W(M|C \text{ und } T) \times W(T|C) \\ &\approx 4,5\% \times W(T|C) \end{aligned}$$

- ▶ einfache Formel zur Korrektur der Verzerrung durch selektives Testen
- ▶ gute, aber selektions-beinflusste Daten werden genutzt
- ▶ aber wie hoch ist  $W(T|C)$  für DE?  
Keine RKI-Daten, keine WHO Daten...
- ▶ Andere Datenquellen?

# Vorschlag einer alternativen IFR-Schätzung

- ▶ Studie der Medizinischen Universität Innsbruck für Corona-Hotspot Ischgl (25.6.2020)  
<https://www.i-med.ac.at/pr/presse/2020/40.html>
- ▶ 79 Prozent der Ischgl (1.473) nahmen an Studie teil
- ▶ Antikörpertest mit hoher Sensitivität (Infektion wird sicher erkannt) und hoher Spezifität (kaum falsch-positiv Ergebnisse).  
Symptomatisch oder nicht symptomatische.
- ▶ Vergleich mit früheren Testergebnissen (wie RKI-Daten mit Selektionseffekten, da nur symptomatische Patienten getestet)
- ▶ *“Anteil der seropositiv Getesteten liegt [...] etwa sechs Mal höher als die Zahl der zuvor [...] positiv getesteten Personen, die Rate der offiziell gemeldeten Fälle beträgt damit nur 15% der de facto Infizierten.”*
- ▶ d.h.  $W(T|C) \approx 0,15!$

## Annahmen

- ▶ In Ischgl wie in DE keine Überlastung des Gesundheitssystems
- ▶ Vergleichbare ursprüngliche Testpraxis wie in DE (hinreichend Symptome)
- ▶ Hohe Sensitivität und Spezifität des Antikörpertests
- ▶ Vergleichbare Alterstruktur der Infizierten
- ▶  $W(T|C) \approx 0,15$  auf DE übertragbar

$$\begin{aligned} IFR = W(M|C) &\approx W(M|C \text{ und } T) \times W(T|C) \\ &\approx 4,5\% \times 15\% \\ &\approx 0,68\% \end{aligned}$$

Geschätzte IFR  $\approx 7$  von 1000 d.h. 7 mal so tödlich wie die Grippe

# Ist die Approximation $IFR \approx 0,68$ plausibel?

Nature 582, 467 – 468 (16. Juni 2020) fasst weltweite Evidenz zusammen:

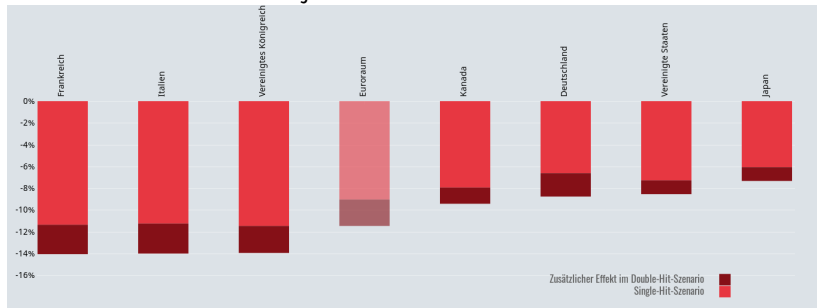
- ▶ *“Public-health researchers use the infection fatality rate to gauge how to respond to a new disease, but it’s tricky to calculate.”*
- ▶ *“Several researchers [...] find it interesting that a growing number of studies from different regions have estimated IFRs in the range of 0.5 - 1%”*
- ▶ *“One of the first studies to account for the effect of age [...] based on seroprevalence data from Geneva, Switzerland, estimates an IFR of 0.6% for the total population. This study is fantastic. It’s precisely what should be done with all of the serological data”*

<https://www.nature.com/articles/d41586-020-01738-2?fbclid=IwAR2jEimZAtoS3KUtmOvbeZ9W5TVBcYsZSbD1jPdPQKQFwdBU549m1lePmqo>



Gravierende makroökonomische Konsequenzen der Maßnahmen zur Eindämmung der Pandemie (siehe Vorlesung Prof. Kohler)

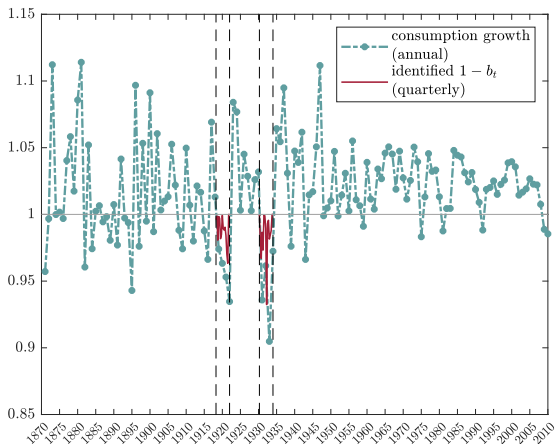
OECD Projektion des BIP-Wachstums G7



Quelle: <http://www.oecd.org/wirtschaftsausblick/juni-2020/>

# USA: 2 Disaster in 130 Jahren, "milder Verlauf" seit 1945

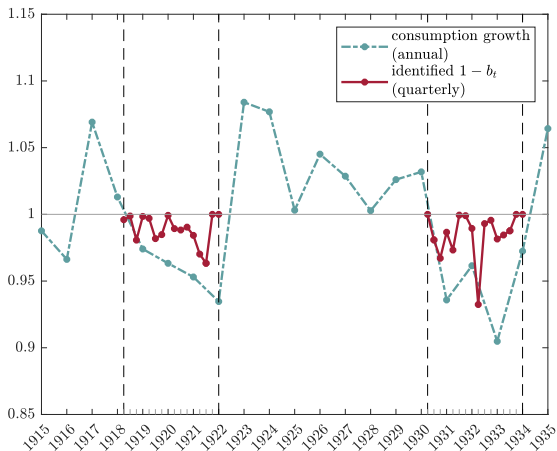
## Brutto-Konsumwachstum in den USA 1870 – 2010



Quelle: Sönksen and Grammig (2020), Barro and Ursúa (2008)

# Zwei Desaster in 130 Jahren

Span. Influenza-Pandemie, 3 Wellen 1918 – 1920, 20 – 50 Mio.  
Tote und Weltwirtschaftskrise



Quelle: Sönksen and Grammig (2020), Barro and Ursúa (2008)

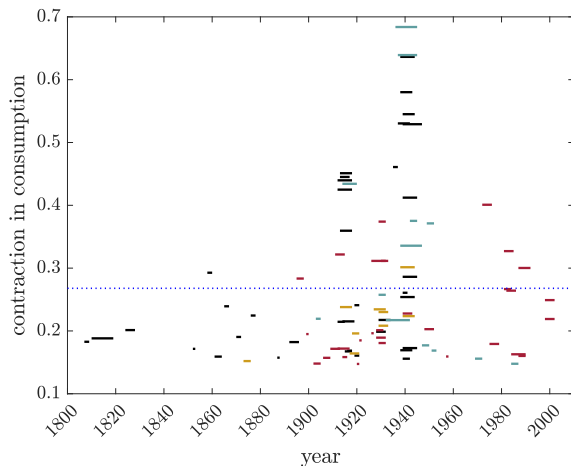
- ▶ Es fällt schwer, Realwirtschaft und Finanzwirtschaft in einem Modell zu verbinden
- ▶ Equity Premium Puzzle / Interest Rate Puzzle
- ▶ Entweder vollkommen unplausibles Anlegerverhalten (Risikoneigung oder Zeitpräferenz)
- ▶ ...und/oder viel zu wenig realwirtschaftliches Risiko, um die Schwankungen auf Finanzmärkten zu erklären

## Sind seltene Desaster das fehlende Puzzleteil?

- ▶ Nur einen “milden Verlauf” beobachtet (bei Corona-Tests gerade umgekehrt)
- ▶ Anleger haben Erwartungen über weniger milde Verläufe berücksichtigt:  
Atomares Wettrüsten (Beginn 1950er), Berlin-Krise, Kuba-Krise (1960er), Kalter Krieg (1960-1990er) Tschernobyl (1986), Sowietunion zerfällt (1990er),...
- ▶ Ökonomische Modelle auf Basis von selektierten Daten (“milder” Verlauf) kalibriert
- ▶ Seltene-Desaster-Hypothese von Barro (2006)
- ▶ Wie kann man das empirisch überprüfen? Es gibt ja nur den beobachteten “milden” Verlauf?

# “Ischgl-Strategie”: Verbindung von Daten!

## 87 Disaster-Ereignisse



Quelle: Sönksen and Grammig (2020)

## “Ischgl-Strategie”: Verbindung disparater Daten!

- ▶ In Sönksen and Grammig (2020) wenden wir die “Ischgl-Strategie” an
- ▶ Umformulierung einer Basisgleichung (oben Totale Wahrscheinlichkeit, hier grundlegende Bewertungsgleichung)
- ▶ Verzerrungskorrektur ähnlich wie oben
- ▶ Plausible Modell-Implicationen: Realwirtschaft und Finanzmärkte können verbunden werden
- ▶ Seltene Desaster als fehlendes Puzzleteil
- ▶ Anleger (die “Märkte”) glauben, dass die Welt risikoreicher ist als die “milden” Verläufe es andeuten, preisen seltene Desaster ein

- ▶ Ökonomische Modelle sollten die Mgl. seltener Desaster (z.B. Pandemie) berücksichtigen (tun sie aber oft nicht)
- ▶ Insbesondere, wenn sie zur Regulierung/Risikomanagement eingesetzt werden. Versicherungsaspekt
- ▶ Regulatorische Implikationen: Eigenkapitalausstattung von Banken
- ▶ Gesundheitssystem
- ▶ Datenselektion und -verzerrung bei der Interpretation von Daten mitdenken
- ▶ Verbindung disparater Daten als Möglichkeit, Selektionsverzerrungen zu korrigieren



Vielen Dank für Ihre Aufmerksamkeit  
Bleiben Sie gesund

Vielen Dank für Ihre Aufmerksamkeit  
Bleiben Sie gesund  
und auf der Hut vor verzerrten  
Stichproben!

- Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, 121:823–866.
- Barro, R. J. and Ursúa, J. F. (2008). Macroeconomic Crises since 1870. *Brookings Papers on Economic Activity*, 39(1):255–335.
- Sönksen, J. and Grammig, J. (2020). Empirical asset pricing with multi-period disasters: A simulation-based approach. European Finance Association, Warsaw 2018 and Econometric Society European Meeting, Cologne 2018. Revise & resubmit *Journal of Econometrics*.