

Wirtschaftswissenschaftliche Methoden: Deskriptive Statistik – Tipps & Tücken

Termin 5

Jakob Kapeller

University of Duisburg-Essen
Institute for Socio-Economics &

Johannes Kepler University Linz

Institute for Comprehensive Analysis of the Economy (ICAE)

Editor: *Heterodox Economics Newsletter*

www.jakob-kapeller.org | www.uni-due.de | www.heterodoxnews.com

UNIVERSITÄT
DUISBURG
ESSEN

Open-Minded

ifSO⁷
institute for
socio-economics

Agenda

- „how to make a good chart?“ –
Einige Hinweise zur Gestaltung von Visualisierungen
- „how to make a manipulative chart?“ –
Einige Hinweise zur kritischen Rezeption von Visualisierungen
 - Graphische Manipulationen
 - Statistische Manipulationen
 - Inhaltliche Manipulationen

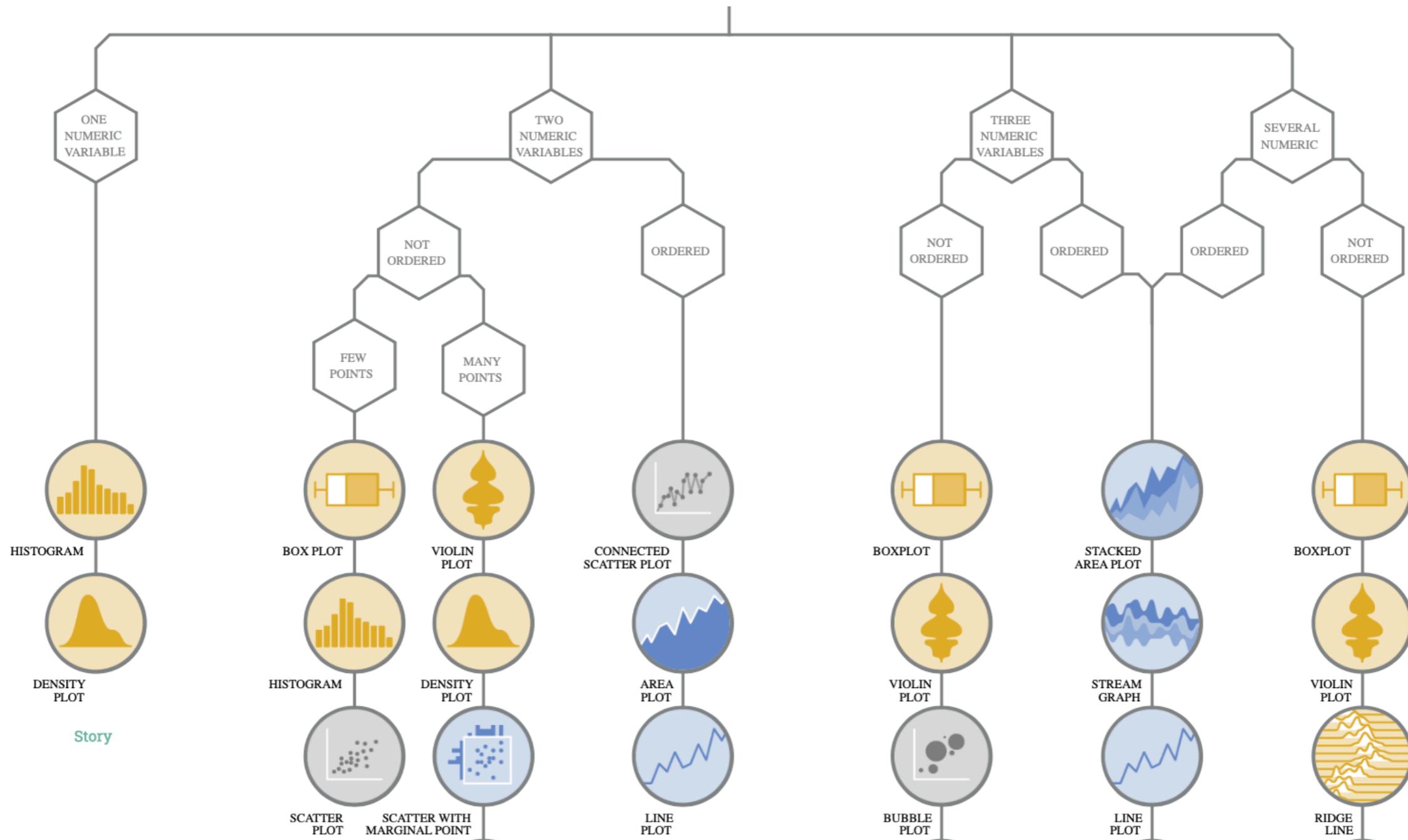
„how to make a good chart“

Wahl geeigneter Visualisierungsformate

- Visualisierungen sollten...
 - ... zur interessierenden Fragestellung passen.
 - ... die Daten in verständlicher Weise zusammenfassen.
 - ... auf die grundsätzliche Struktur der Daten Rücksicht nehmen.
 - ... keine manipulativen Elemente enthalten.
 - ... weitgehend selbsterklärend und
 - ... nicht hässlich sein.
- Tipp zur Findung passender Visualisierungen: <https://www.data-to-viz.com>

Passende Visualisierungen finden: Webtipp

<https://www.data-to-viz.com>



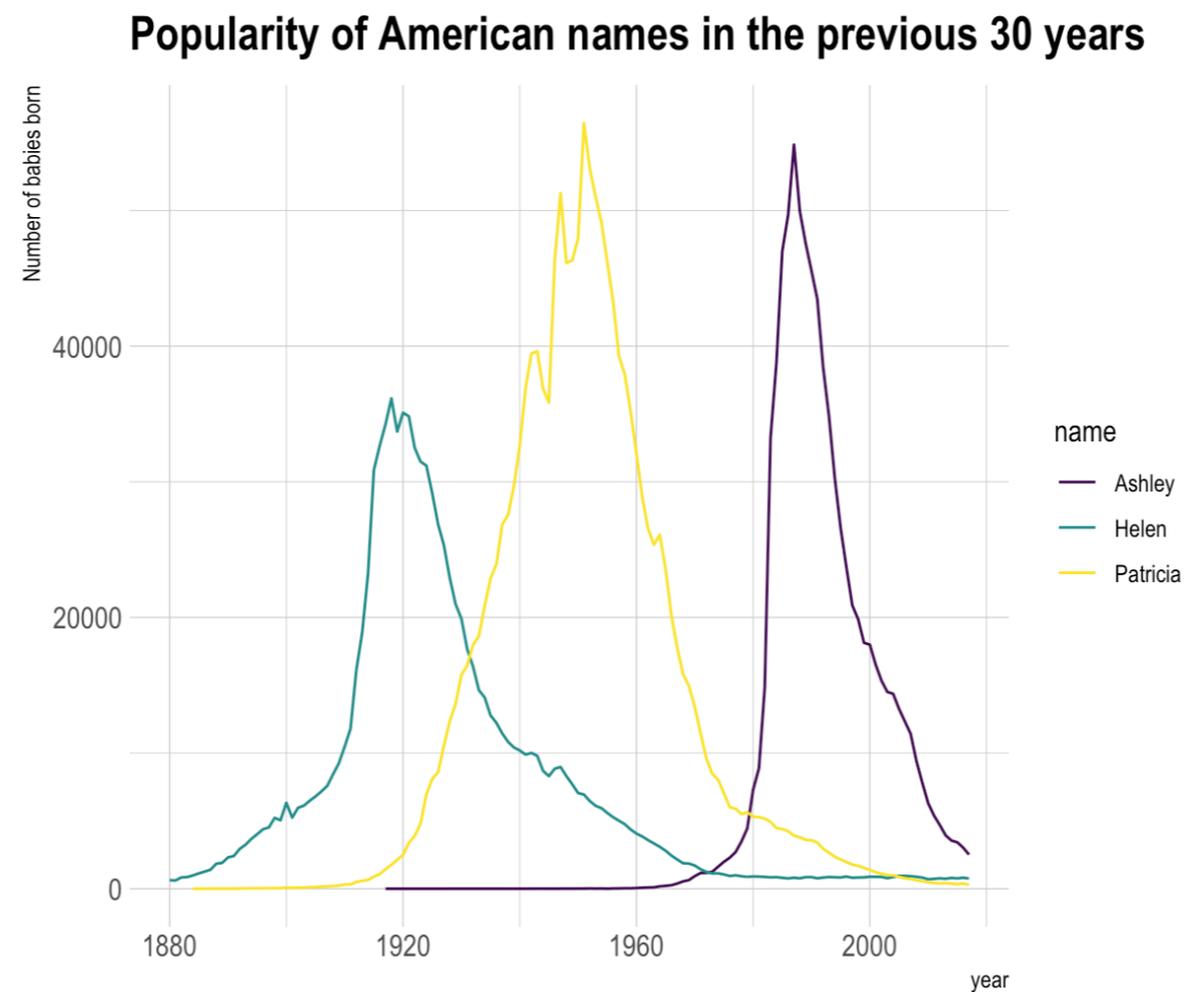
Wahl geeigneter Visualisierungsformate

Vier Klassiker

- Line Chart
 - Scatterplot
 - Histogramm
 - Bubble Chart
-
- Bei allen Visualisierungen ist es zentral...
 - ... der Visualisierung einen Titel zu geben.
 - ... die Achsen zu beschriften.
 - ... zu prüfen, ob die (an den Achsen) angegebenen Einheiten nachvollziehbar und sinnvoll sind.

Ein Klassiker (1): Die Liniengraphik

- Gut zum Beschreiben von Dynamiken (insbesondere Zeitreihen)
 - Grundsätzlich immer verwendbar, wenn die x-Achse die Daten adäquat ordnet (z.B. Verteilungsgraphiken).



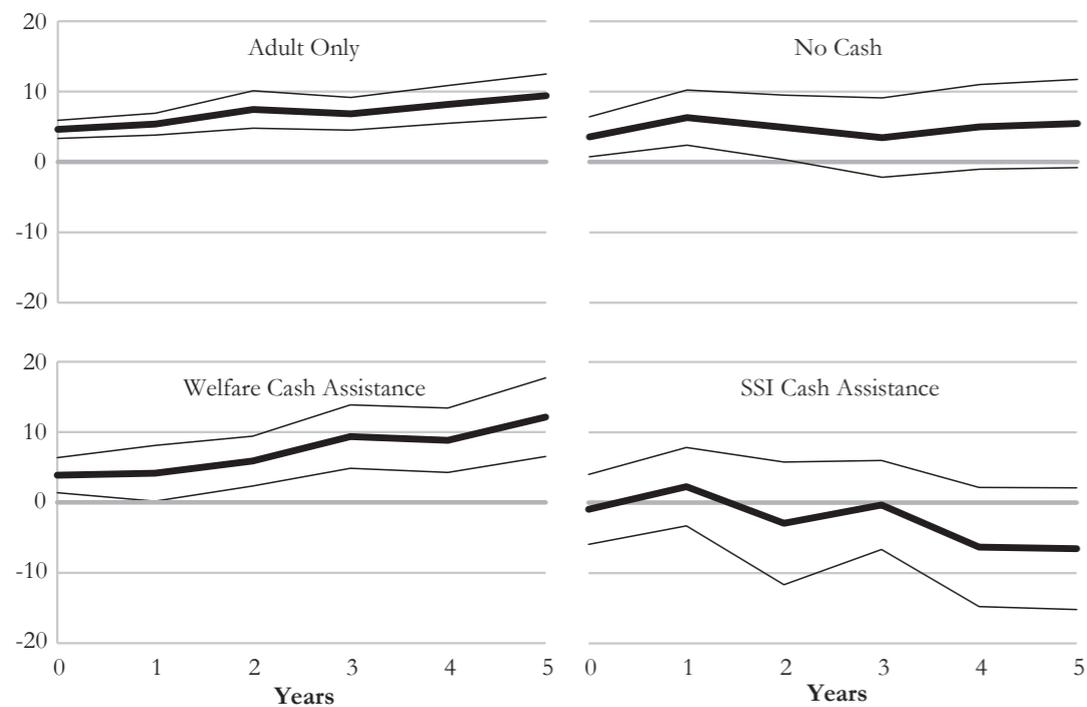
Good line charts

- Add-ons: Plotten von Konfidenzintervallen (links) und Ausweisen der Datenpunkte (rechts)

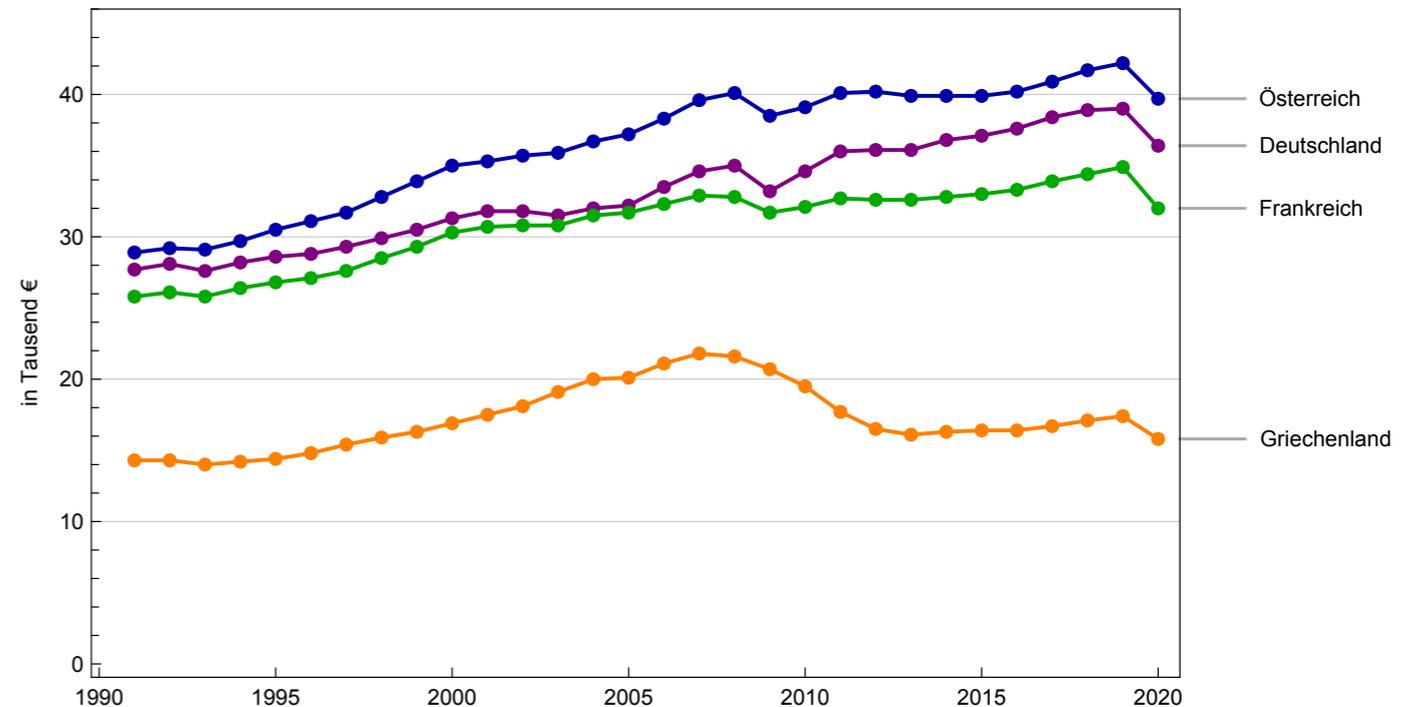
Figure 1B

A Revised Line Chart

Implied Impulse Response Functions for Different Caseloads
(Percent change)



BIP/Kopf in vier europäischen Ländern
(real, zu Preisen von 2015)

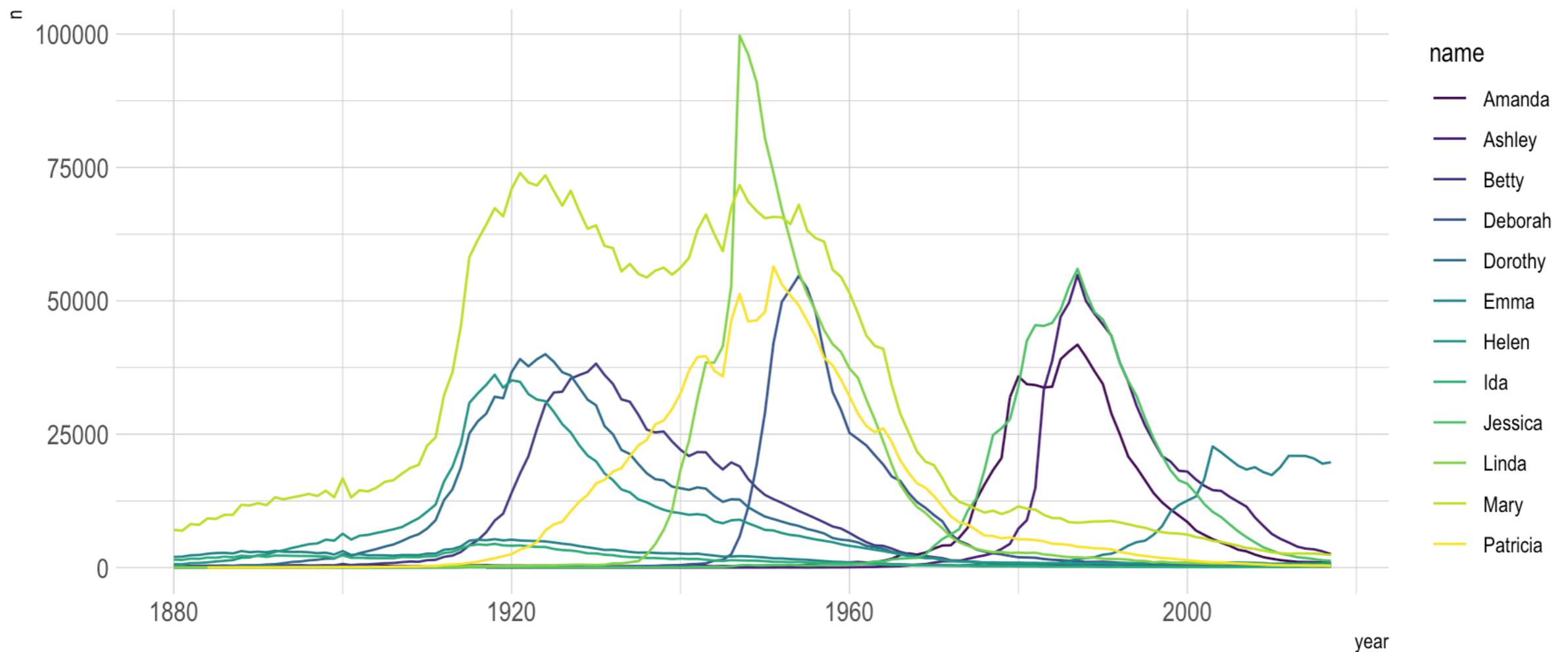


Gefahrenquellen beim Linechart

Der „Spaghetti-Chart“

- Die Liniengraphik wird schnell unübersichtlich...

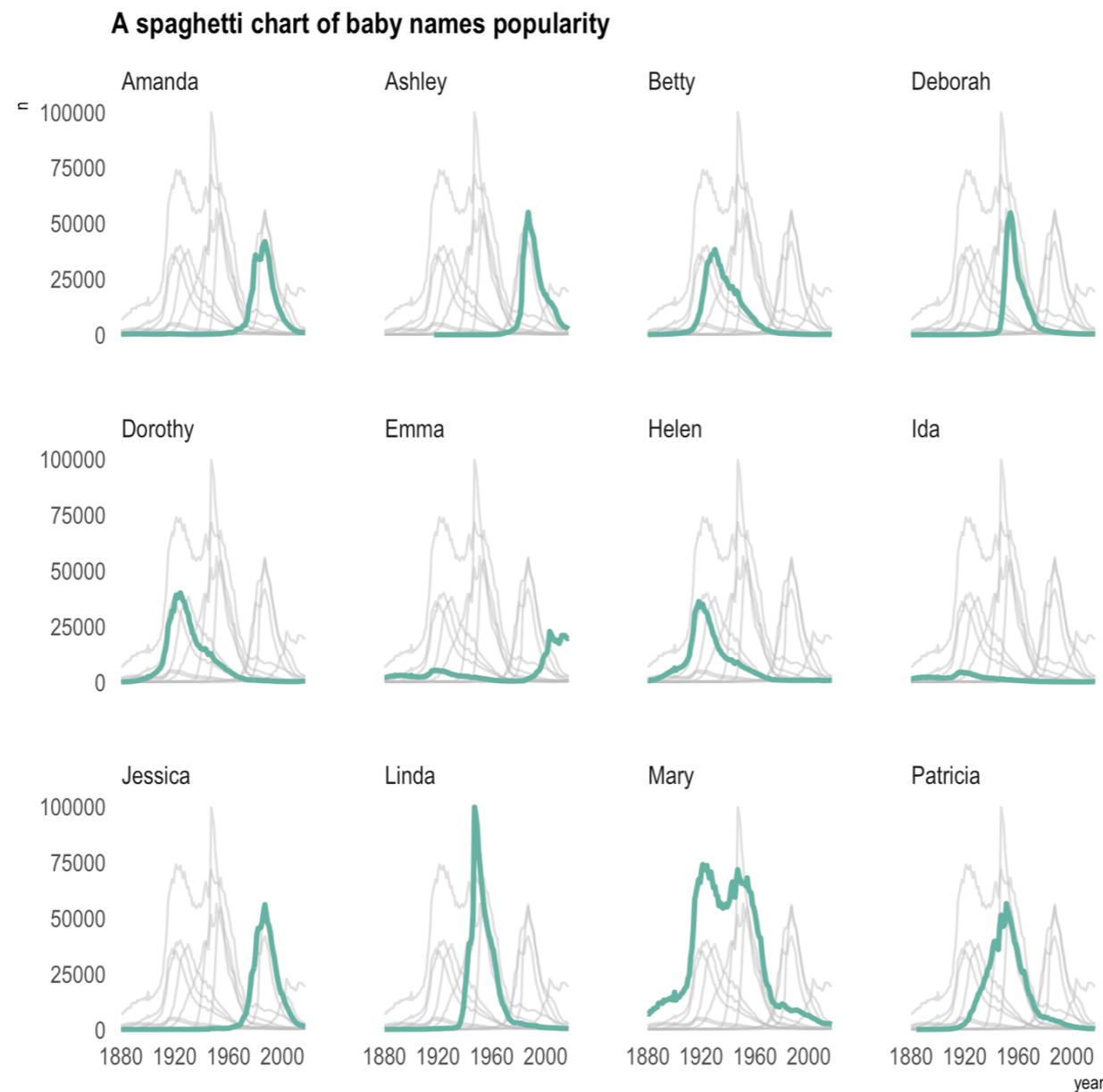
A spaghetti chart of baby names popularity



Gefahrenquellen beim Linechart

Resolving the Spaghetti Chart

- Die Liniengraphik wird schnell unübersichtlich...

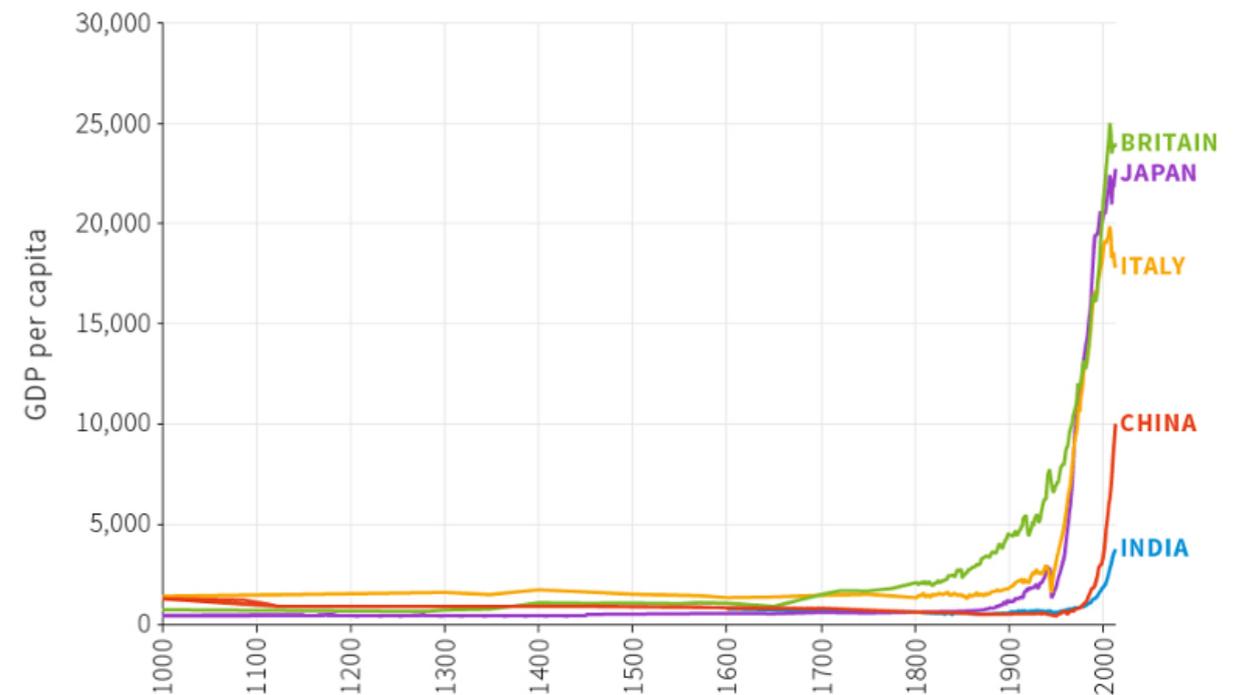
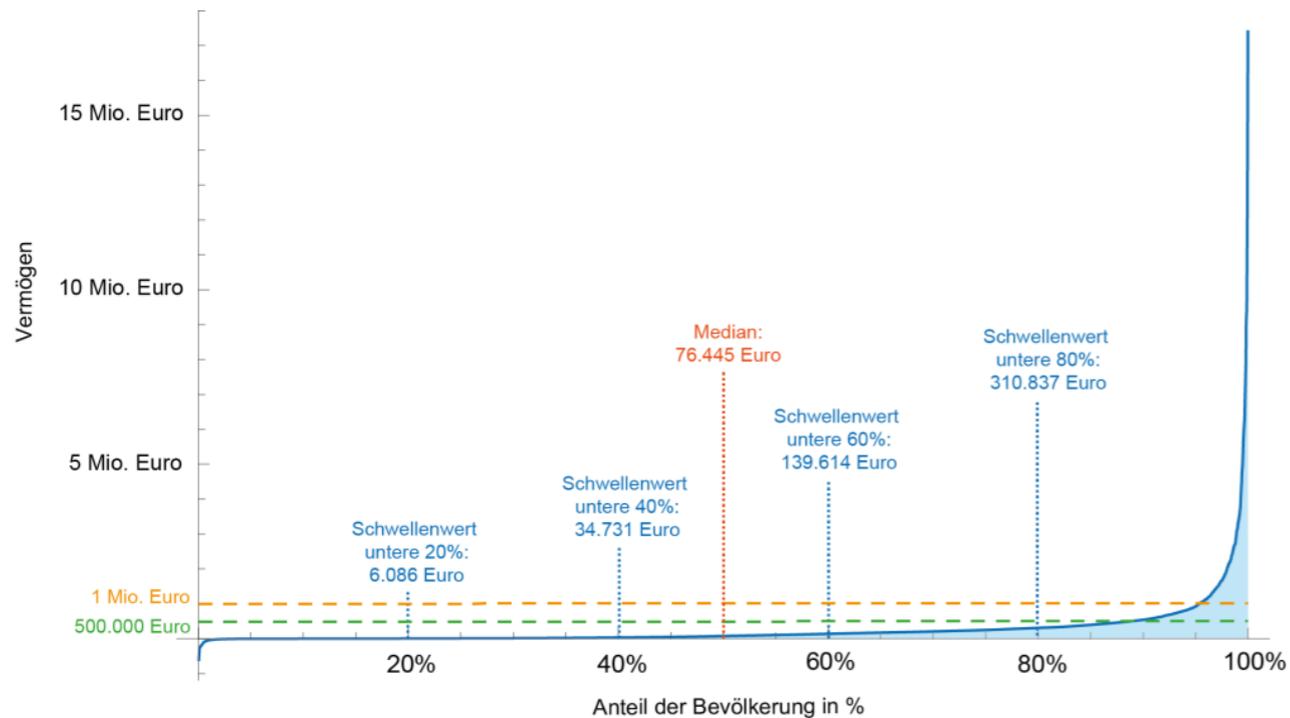


<https://www.data-to-viz.com/caveat/spaghetti.html>

Gefahrenquellen beim Linechart

Skalierungseffekte (1): Bei großer Variation verschwinden kleine Unterschiede

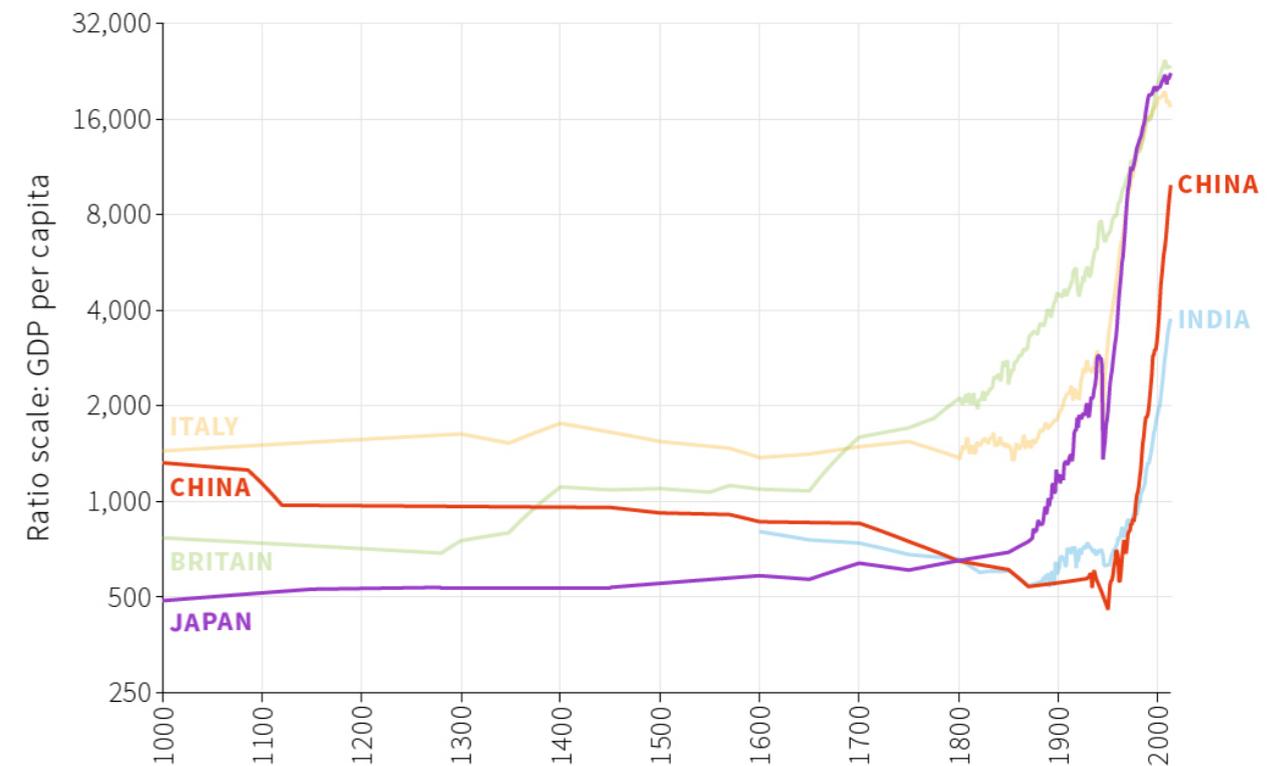
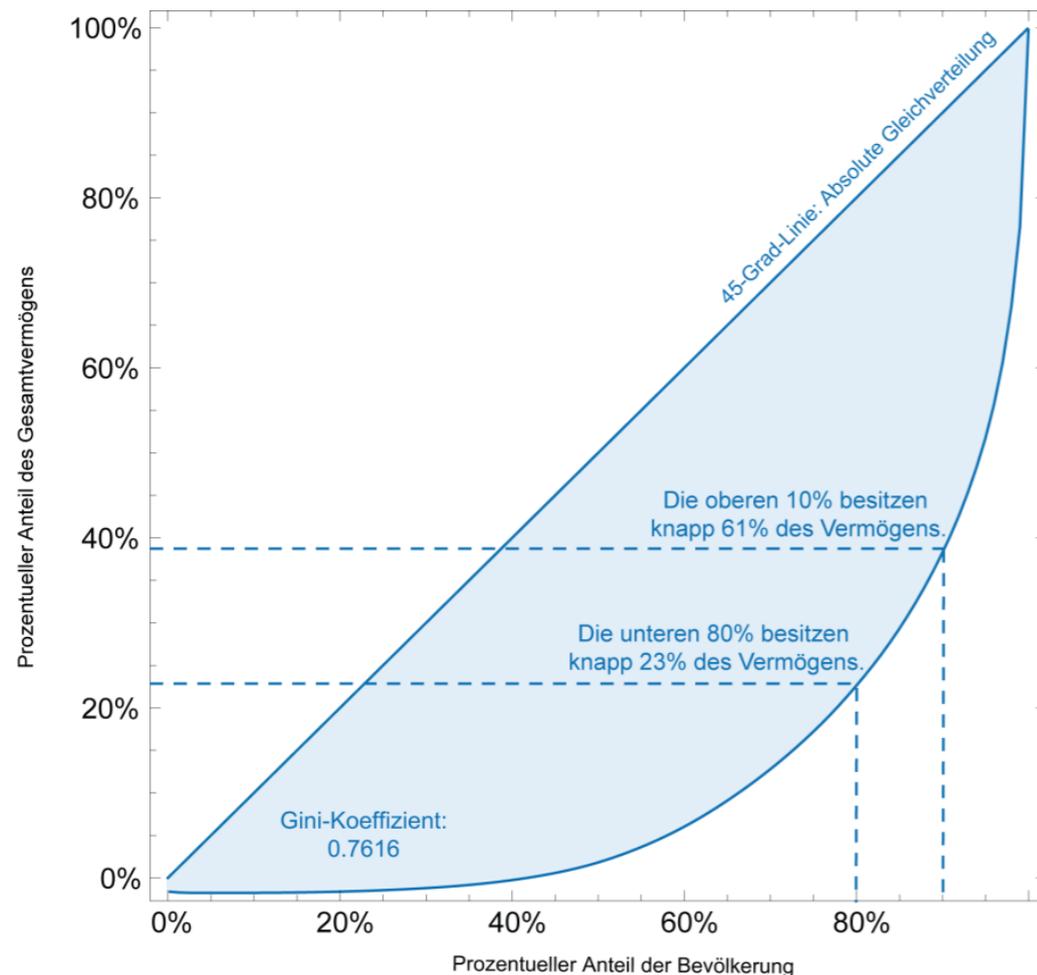
- Fokus auf absolute Änderung prägt Darstellung
 - Kann erwünschte und unerwünschte Effekte gleichzeitig haben.



Gefahrenquellen beim Linechart

Skalierungseffekte (1): Bei großer Variation verschwinden kleine Unterschiede

- Fokus auf absolute Änderung prägt Darstellung
 - Mögliches Komplement / Mögliche Alternative: **Fokus auf relative Änderung**
 - **Beispiele:** Reskalierung nach relativen Vermögensanteilen (links) und Logarithmus (rechts)

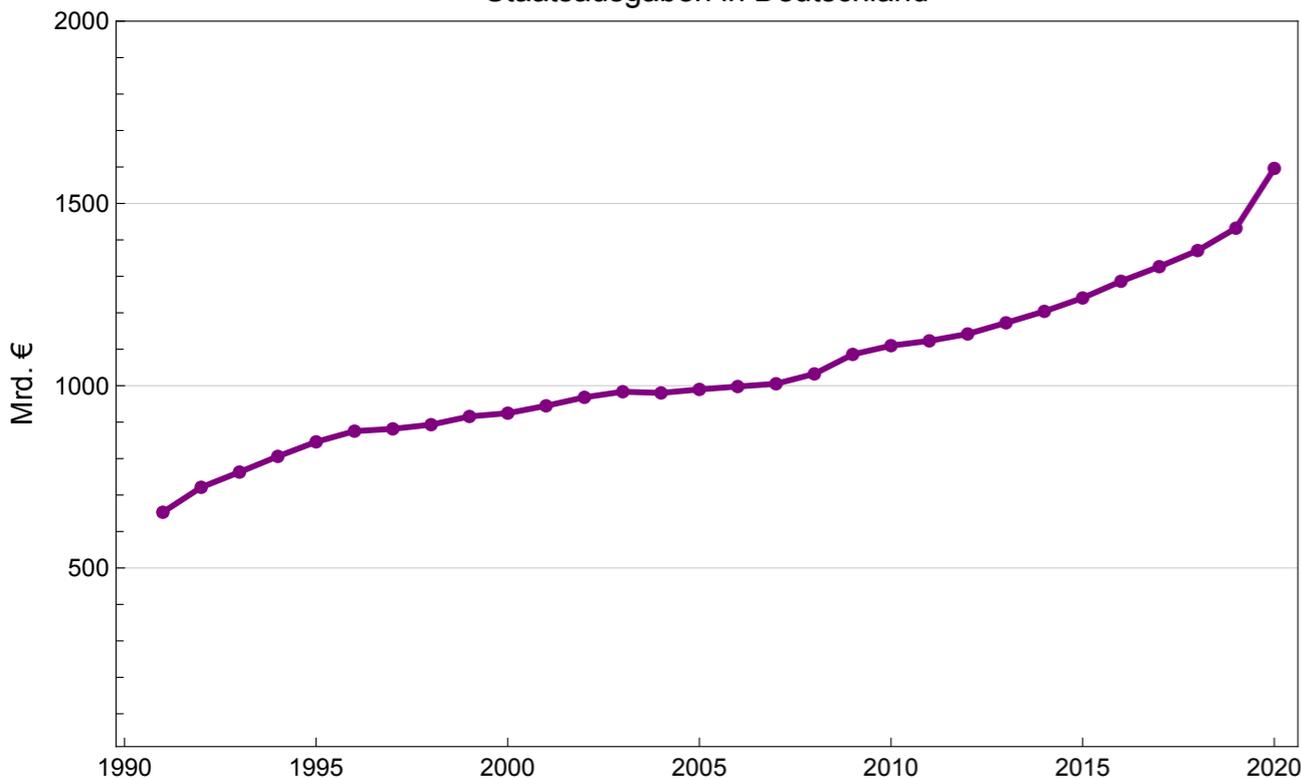


Gefahrenquellen beim Linechart

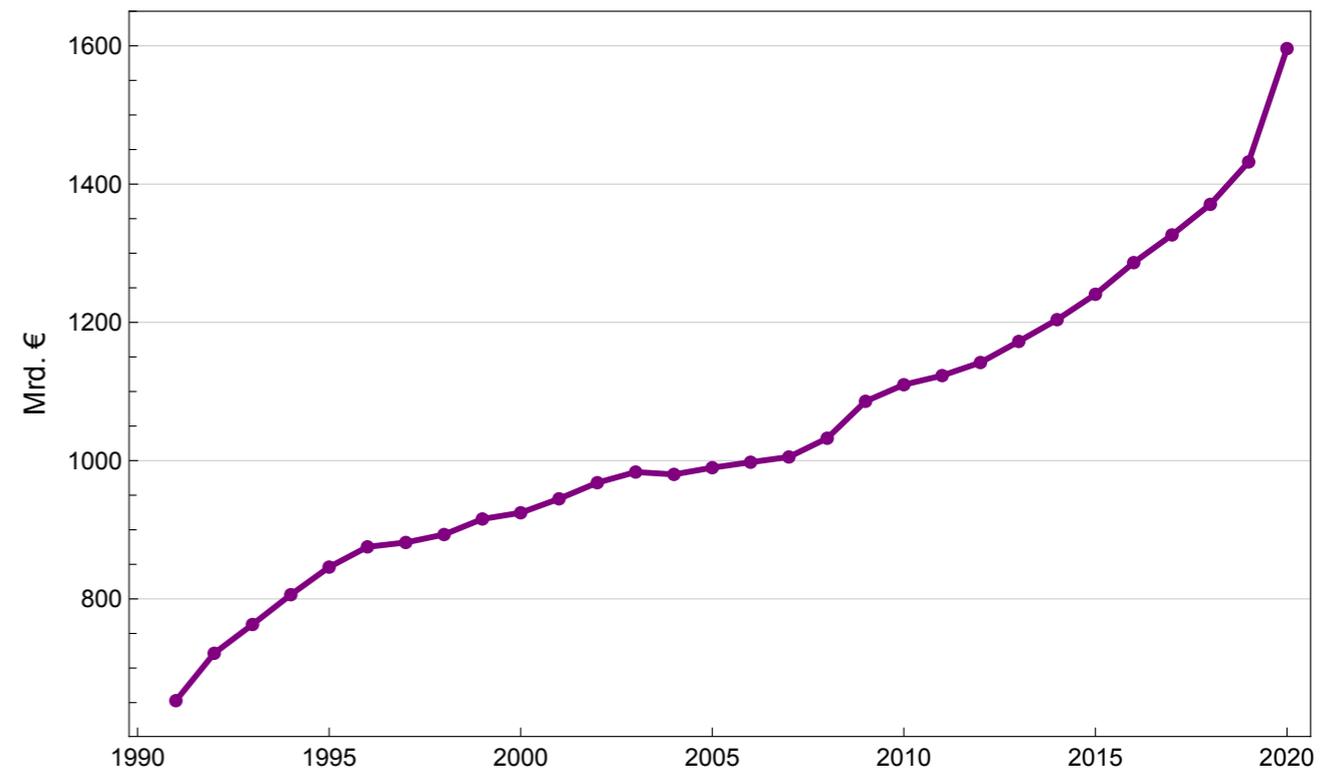
Skalierungseffekte (2): Achsenskalierung beeinflusst optischen Effekt

- **Beispiel:** Wie dramatisch ist der Anstieg der Staatsausgaben?
- Grundsätzliches Problem tritt auch in anderen Darstellungsformen auf.

Staatsausgaben in Deutschland

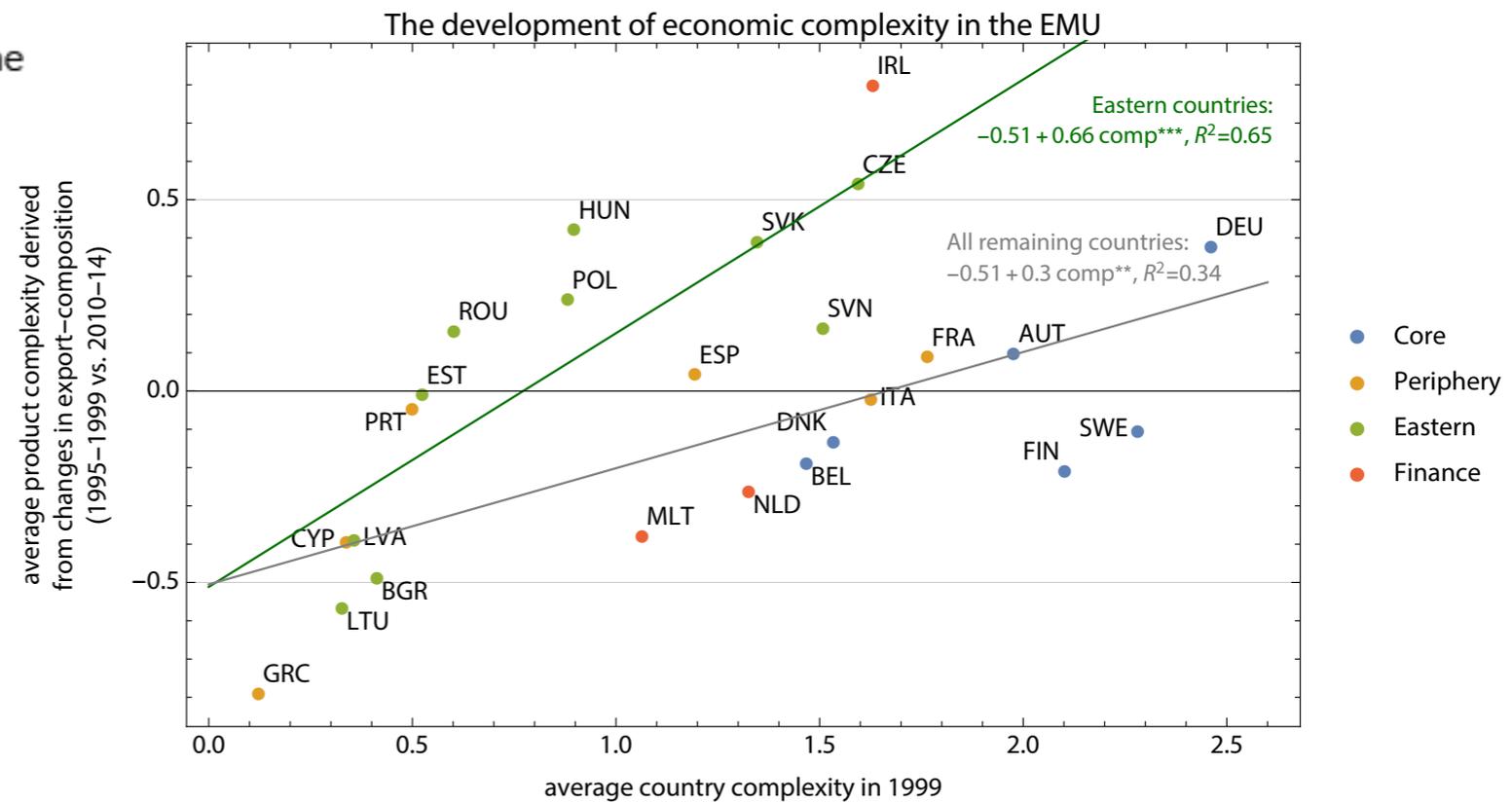


Staatsausgaben in Deutschland



Ein Klassiker (2): Der Scatterplot

- Gut zum Beschreiben von Zusammenhängen zwischen Variablen
 - Kombinierbar mit (unterschiedlichen Varianten von) Regressionslinien (+ error bars).
 - Einmal abstrakt (links), einmal mit Anwendung auf technologische Entwicklungsdynamik („complexity“, rechts).



Gräbner, Claudius; Heimberger, Philipp; Kapeller, Jakob (2019): Wirtschaftliche Polarisierung in Europa. <https://www.fes.de/wirtschaftliche-polarisierung-in-europa>

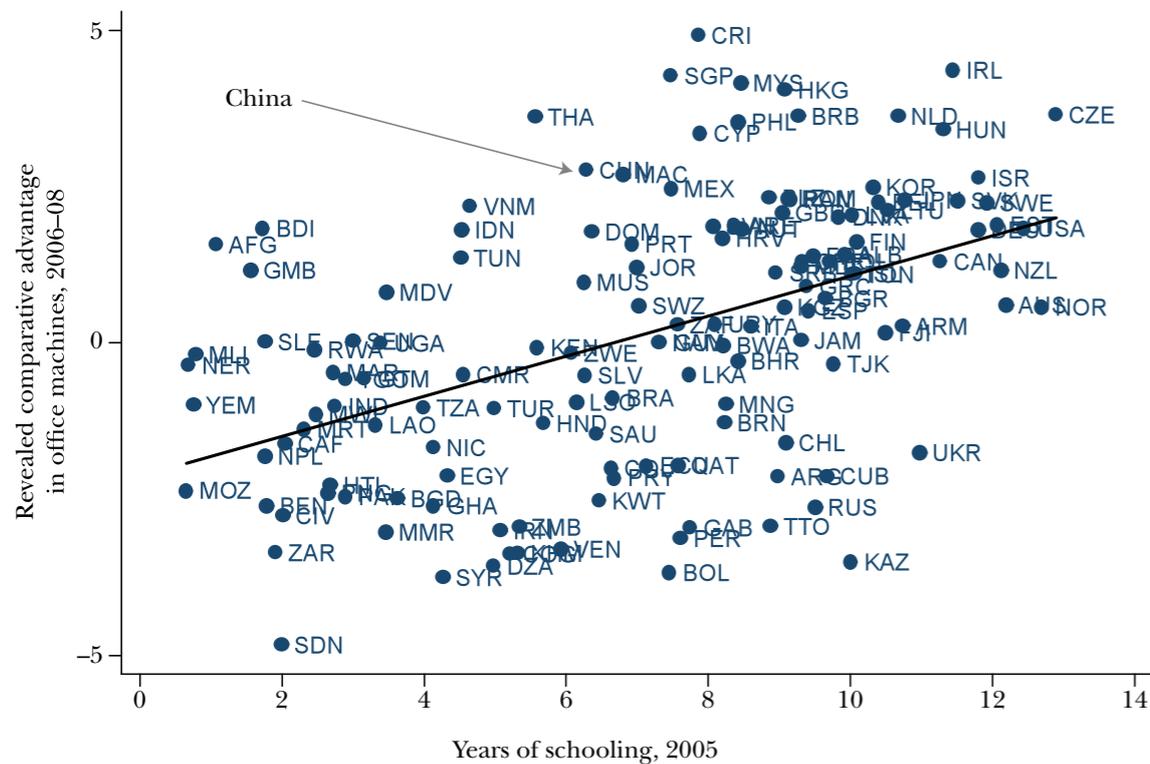
Gefahrenquellen beim Scatterplot

Der „Clutterplot“

- „Clutterplot“: Zu viele Datenpunkte und Beschriftungen – fehlende Übersichtlichkeit.

Figure 2A
A Clutterplot Example

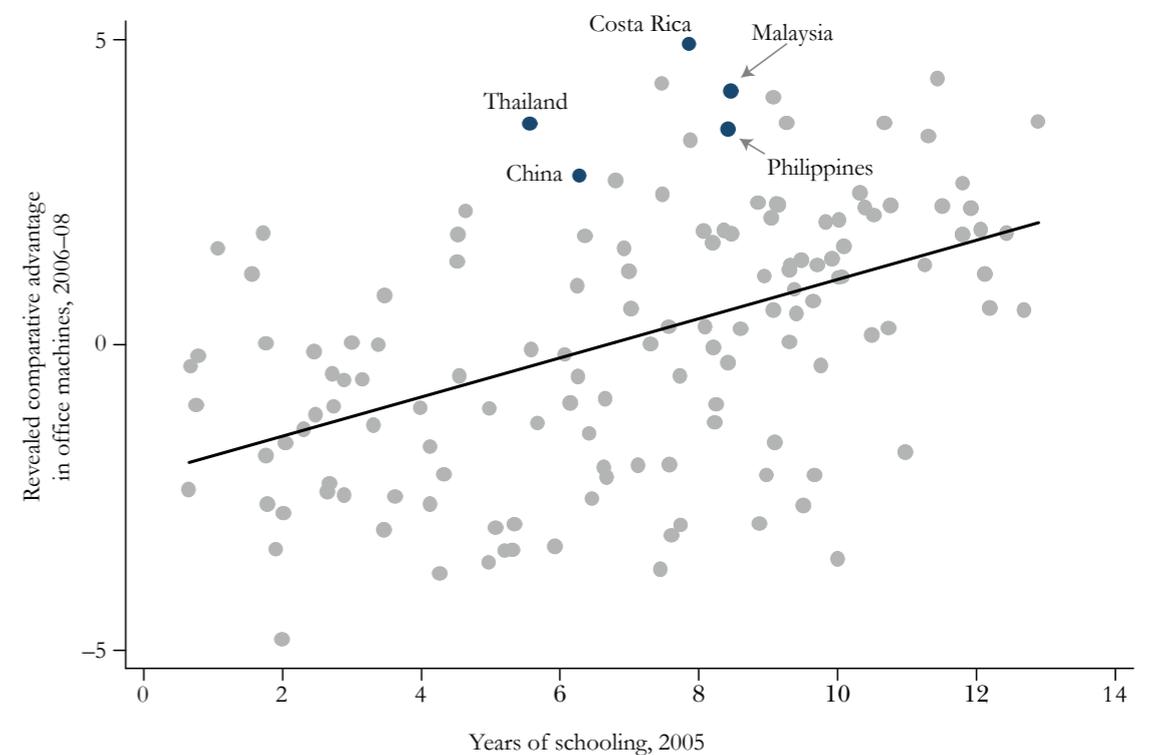
Education and Exports of Office Machines



Source: Hanson (2012).

Figure 2B
Revising the Clutterplot Example

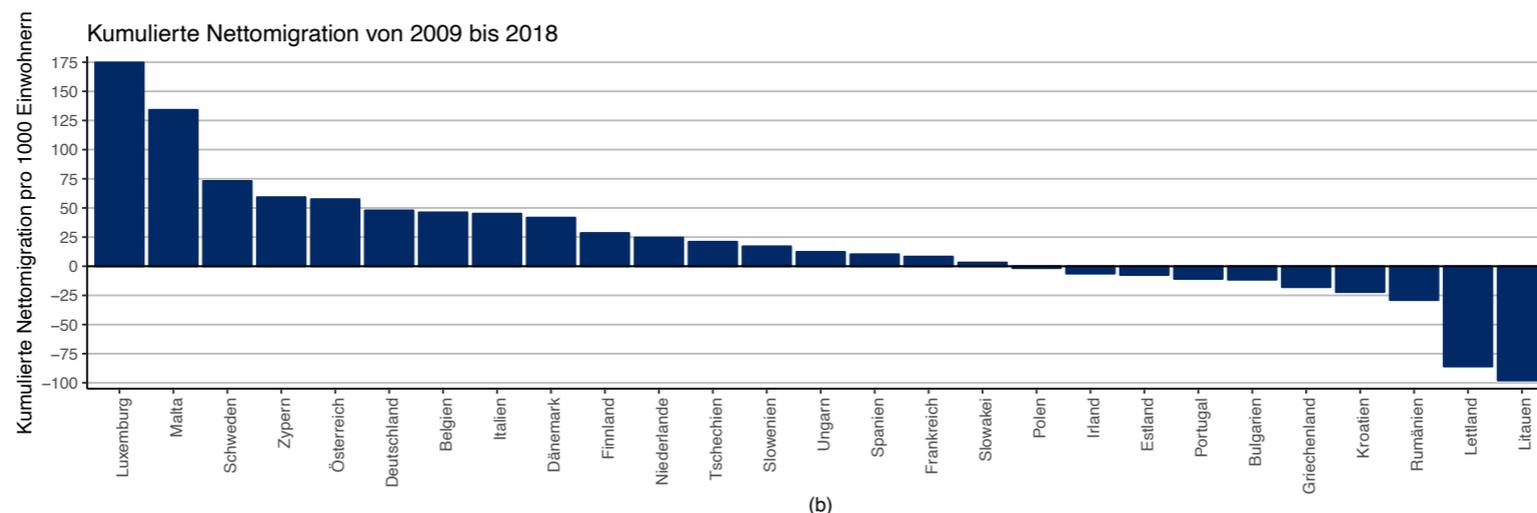
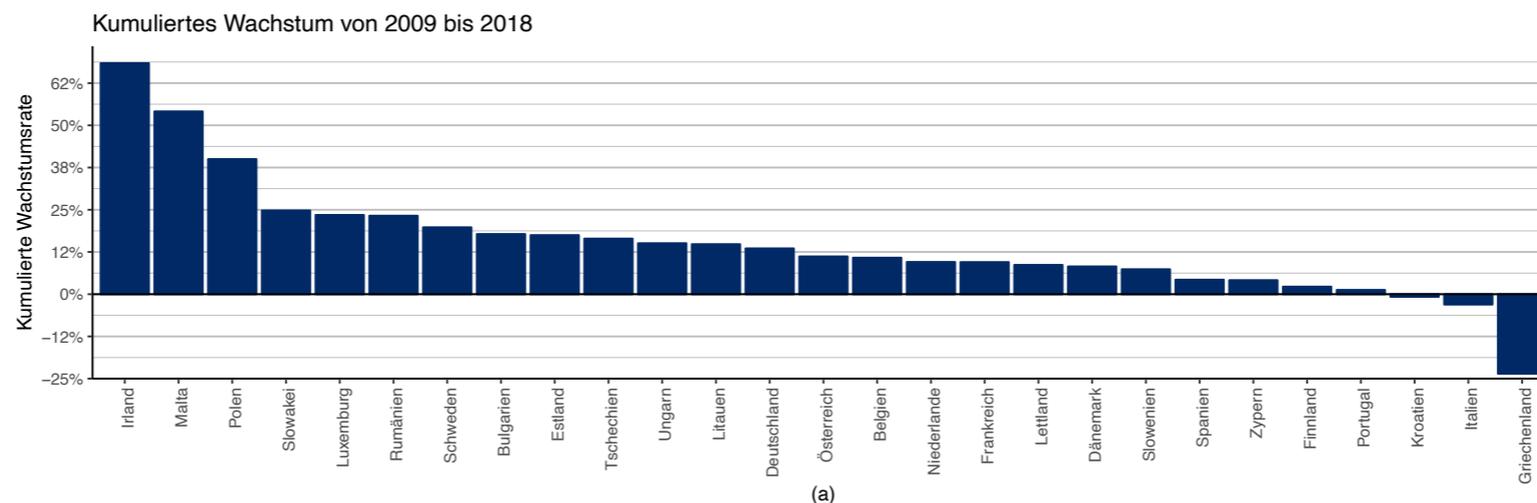
Education and Exports of Office Machines



Schwabish, J. (2014). An Economist's Guide to Visualizing Data Journal of Economic Perspectives 28(1), 209 - 234.

Ein Klassiker (3): Der Barchart (das Histogramm)

- Gut zum Beschreiben der allgemeinen Verteilung einer Variable
 - Große Freiheitsgrade in der Gestaltung: Unterschiedliche Gruppen, „raw“ vs. „stacked“, Säulen vs. Balken etc.

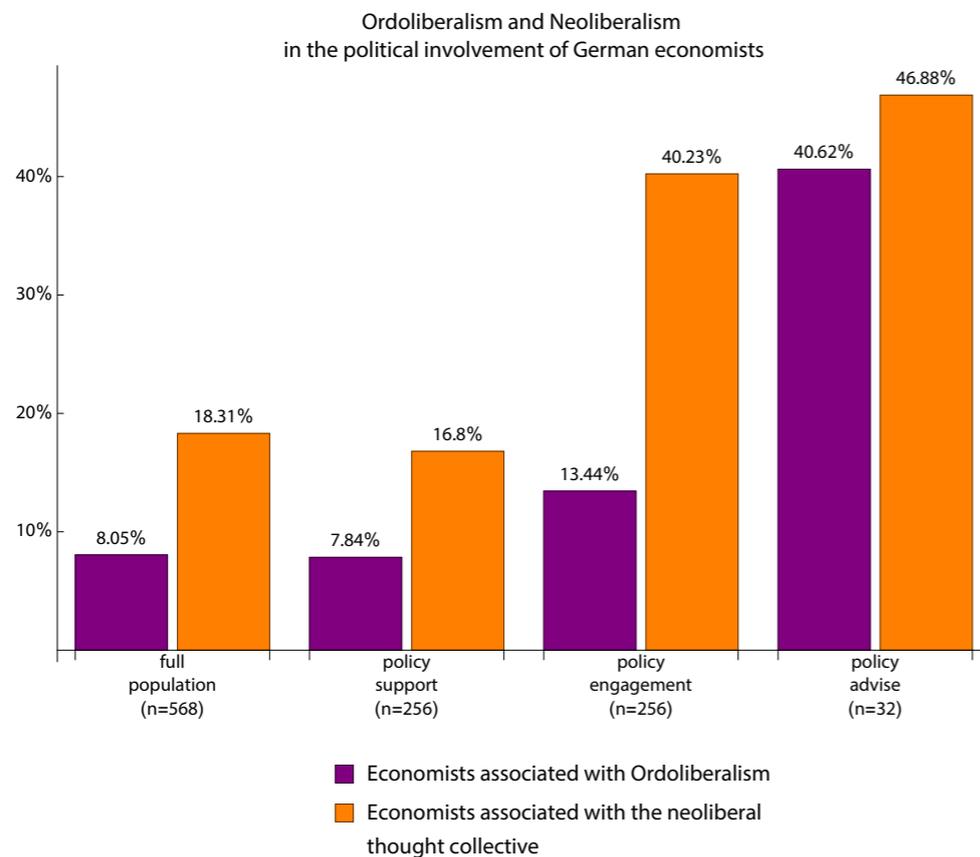


Gräbner, Claudius; Heimberger, Philipp; Kapeller, Jakob (2019): Wirtschaftliche Polarisierung in Europa. <https://www.fes.de/wirtschaftliche-polarisierung-in-europa>

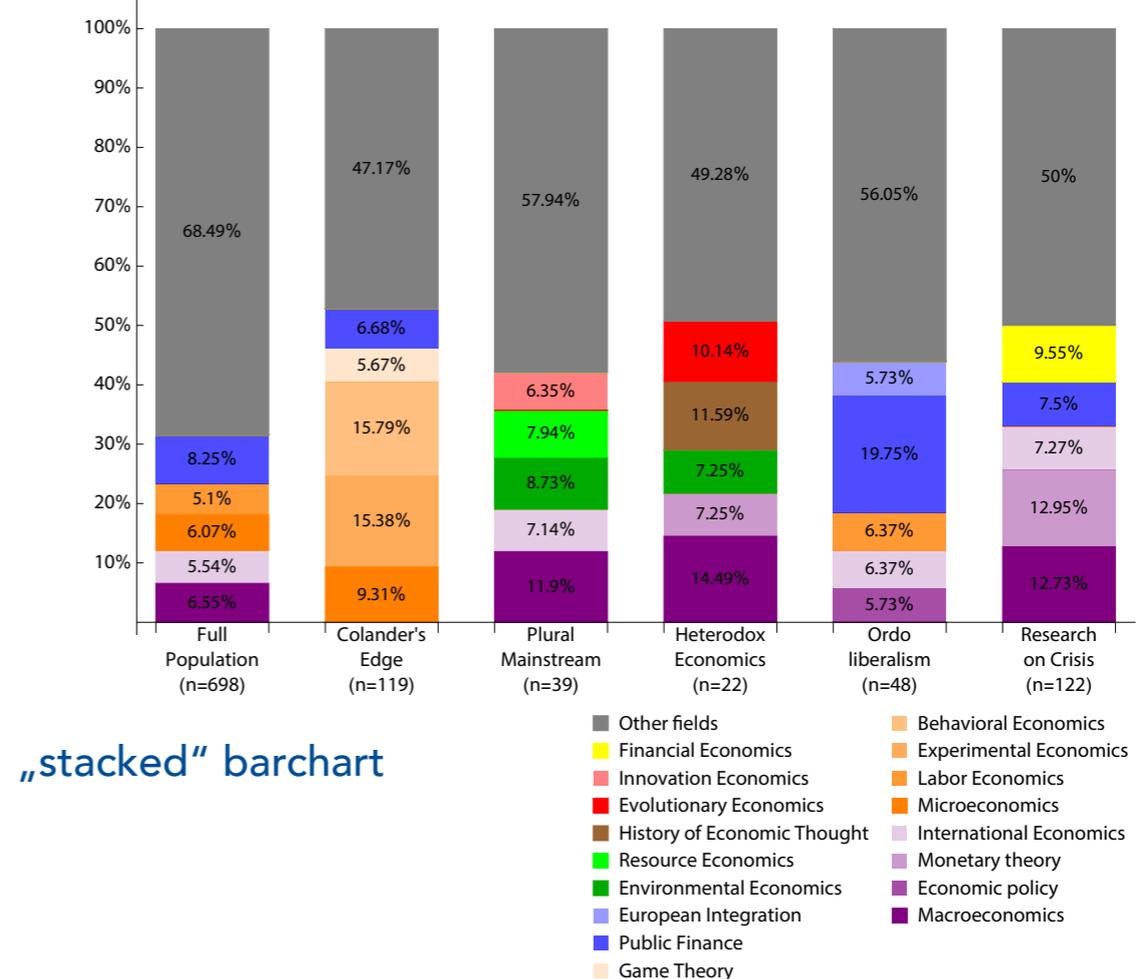
Ein Klassiker (3): Der Barchart (das Histogramm)

- Beispiel: Politische Orientierung und Forschungsinteresse dt. ÖkonomInnen
 - Hier: Unterschiedliche Gruppen, „raw“ vs. „stacked“

Zwei Variablen für vier Gruppen



Research focus and topical orientation of economists in the German-speaking countries

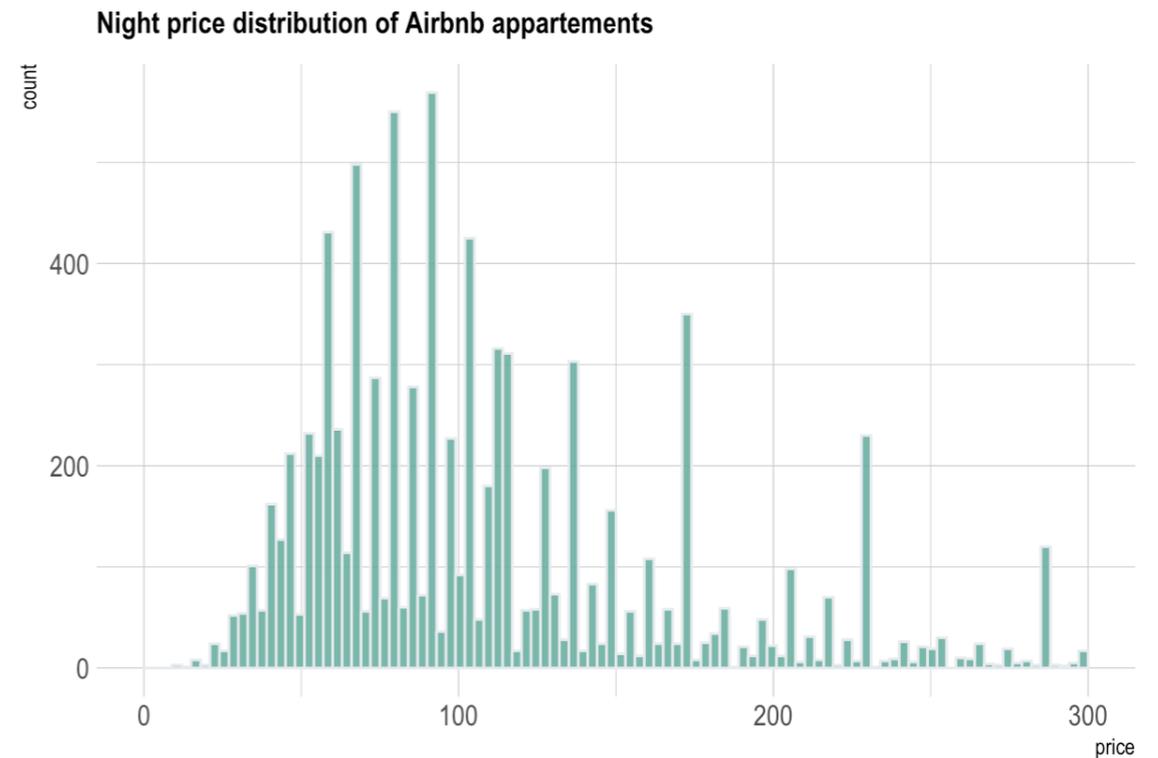
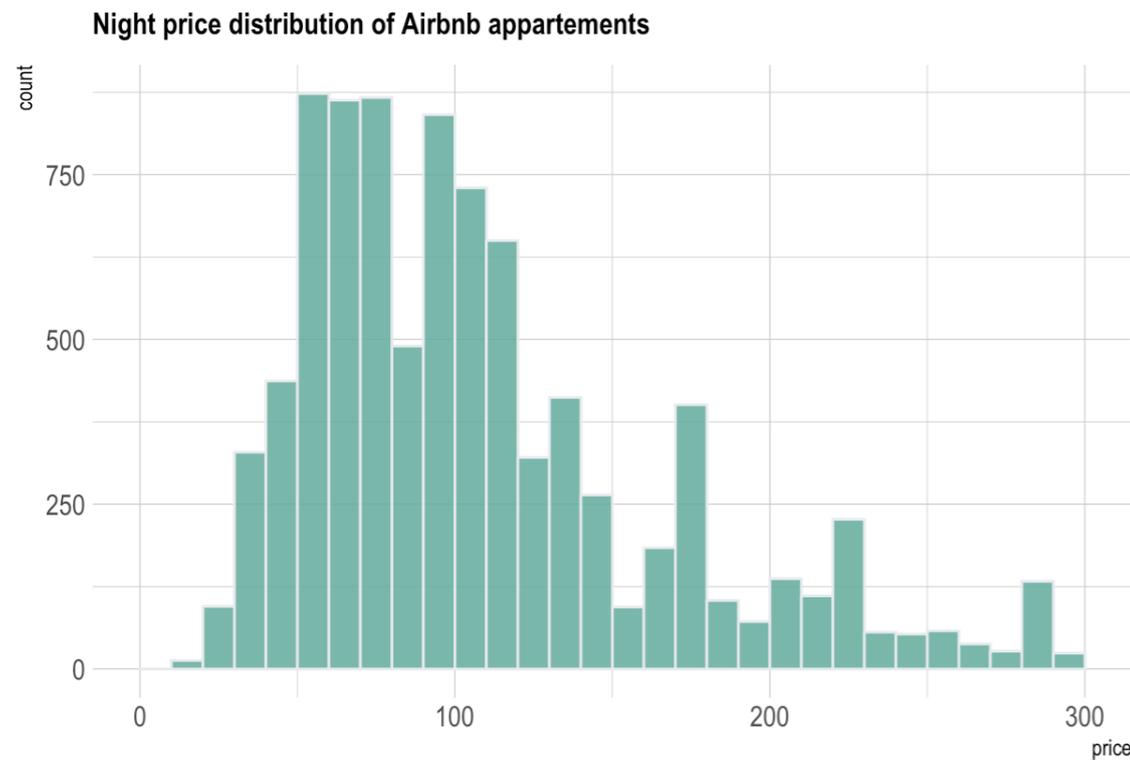


Grimm Christian, Kapeller Jakob and Stephan Pühringer: Paradigms and Policies: The state of economics in the german-speaking countries. ICAE Working Paper #77

Ein Klassiker (3): Der Barchart (das Histogramm)

Tipp: Breite der „bins“ variieren...

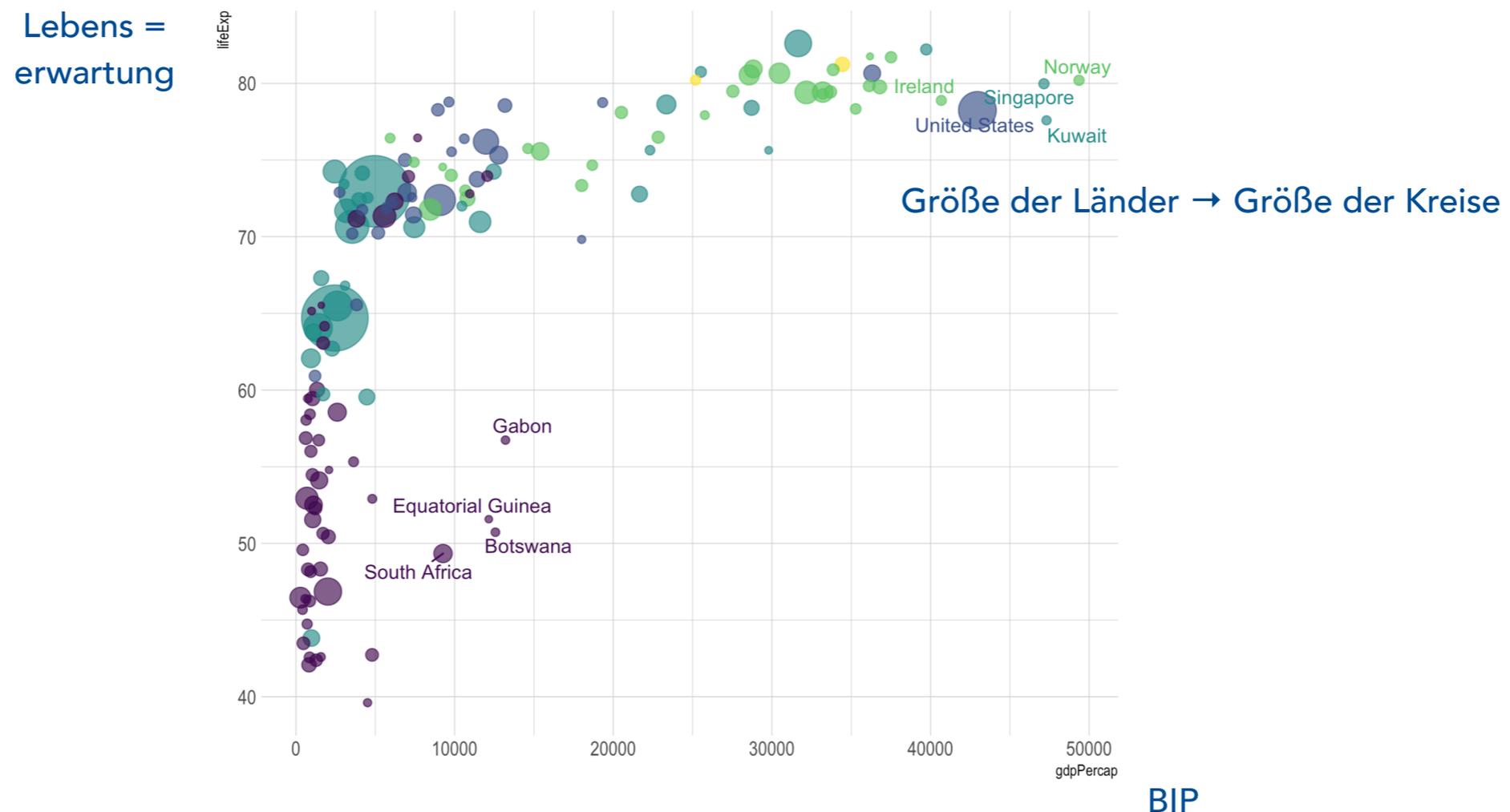
- Beispiel: Mietpreise auf Airbnb



https://www.data-to-viz.com/caveat/bin_size.html

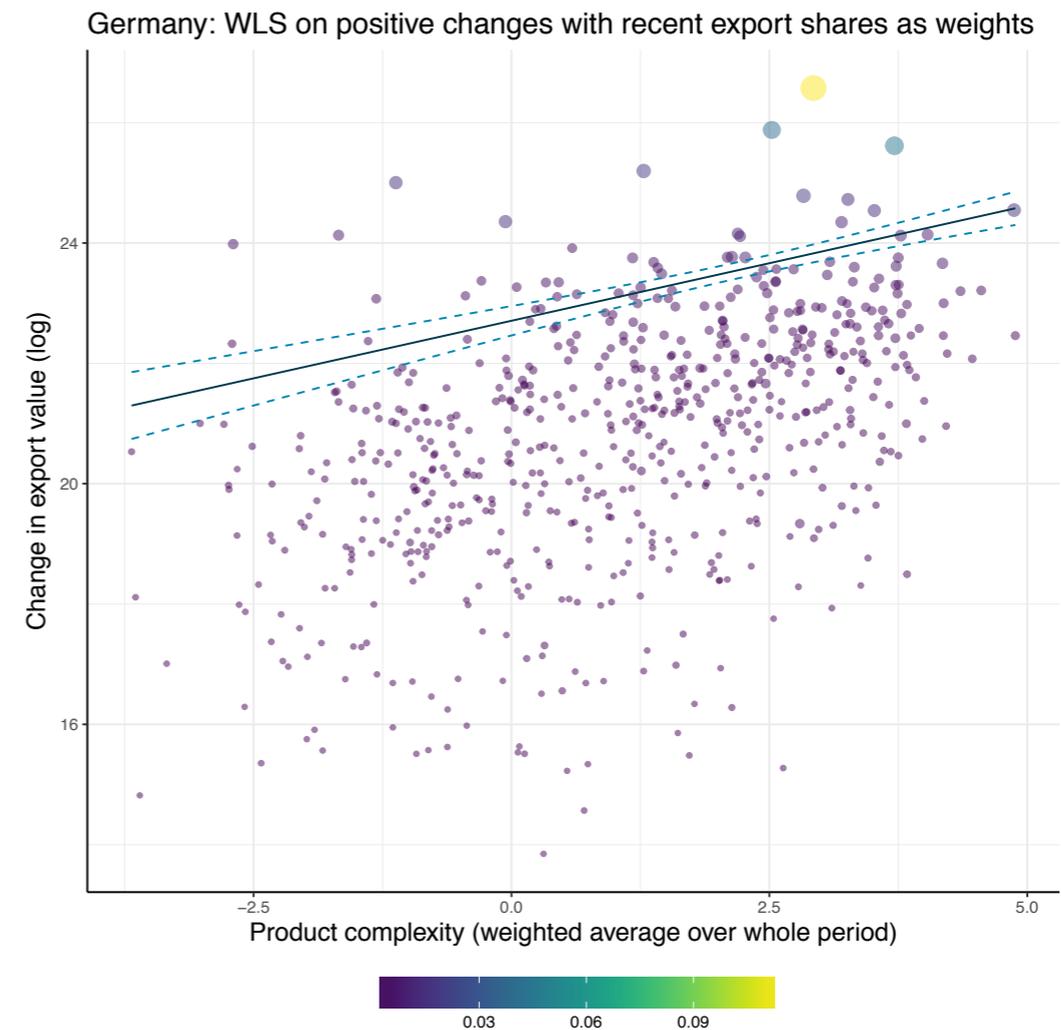
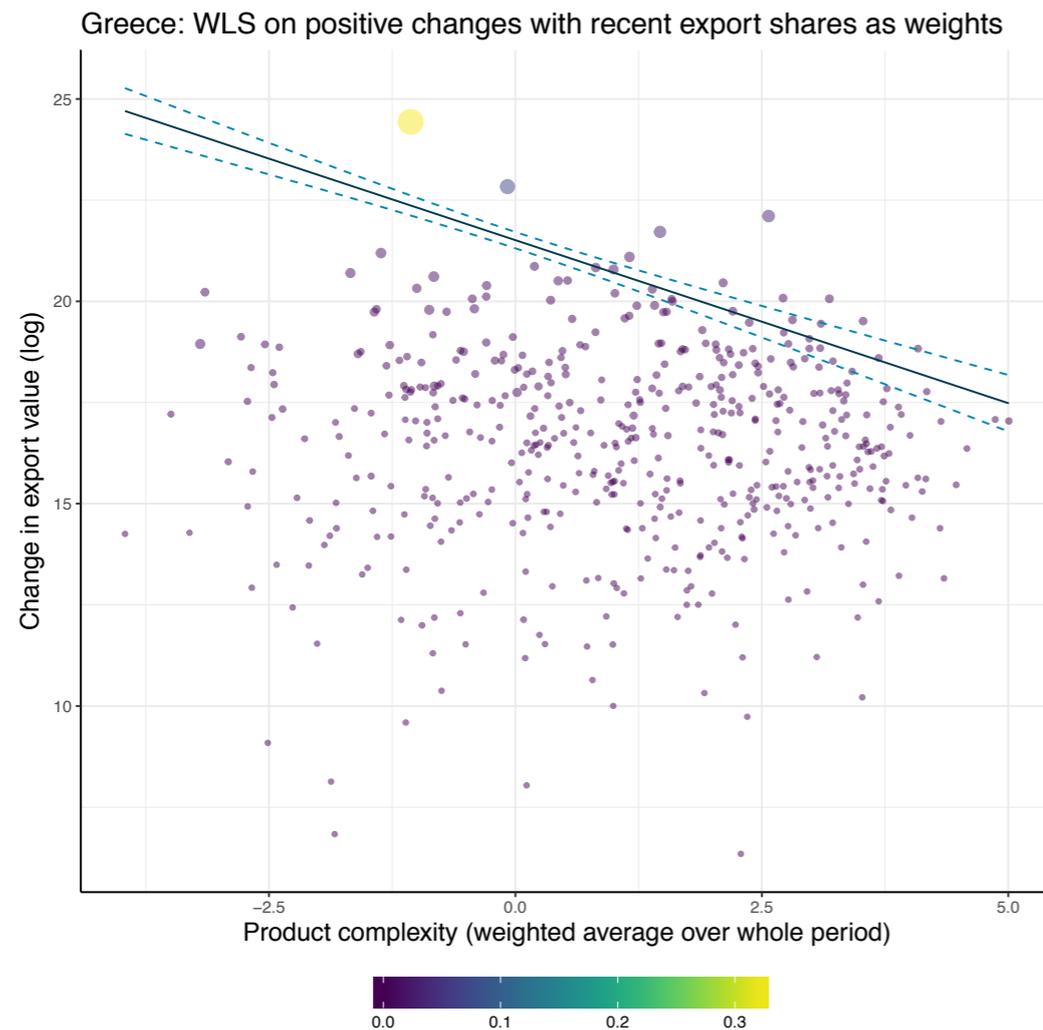
Ein Klassiker (4): Der Bubbleplot

- Immer praktisch, wenn man auf die Heterogenität der Beobachtungsgegenstände eingehen möchte.
 - Kann auch mit anderen Darstellungsformen kombiniert werden... (meist: Scatterplot)



Ein Klassiker (4): Der Bubbleplot

- Sektorales Wachstum und technologische Komplexität
 - Größe der Sektoren als Bubbles – schafft Orientierung. Zusätzlich Farbcode.



Gräbner et al. (2018): Structural Change in times of increasing openness. ICAE WP #76

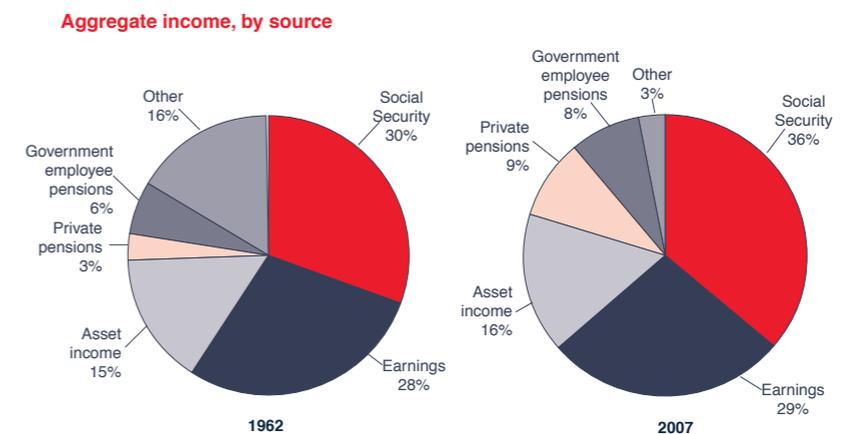
Tortendiagramm: Kein Klassiker im positivem Sinne

„The pie-chart and its alternatives...“

- Allgemein: Tortendiagramme eher vermeiden.
 - **Grund #1:** Winkel sind für menschlichen Verstand schwer zu interpretieren.
 - **Grund #2:** Insbesondere bei komplexeren Sachverhalten (z.B. Vergleichen) sind Tortendiagramme unnötig kompliziert.
- Daher: Alternativen!

Figure 9A
Two Pie Charts for Comparison

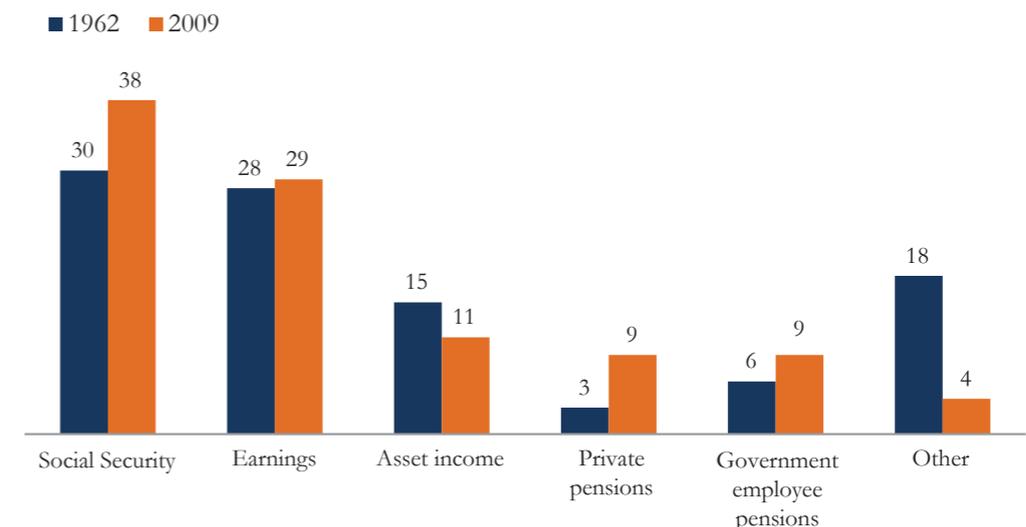
Shares of Aggregate Income, 1962 and 2007



Source: Social Security Administration (2009).

Figure 9B
Alternative to a Pie Chart: A Paired Column Chart

Shares of Aggregate Income, 1962 and 2009 (Percent)



Tortendiagramm: Kein Klassiker im positivem Sinne

„The pie-chart and its alternatives...“

Figure 8A
A Pie Chart, Labeled

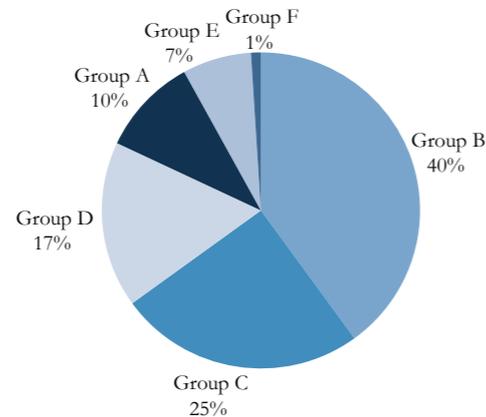


Figure 8B
Pie Chart Alternative: A Bar or Column Chart

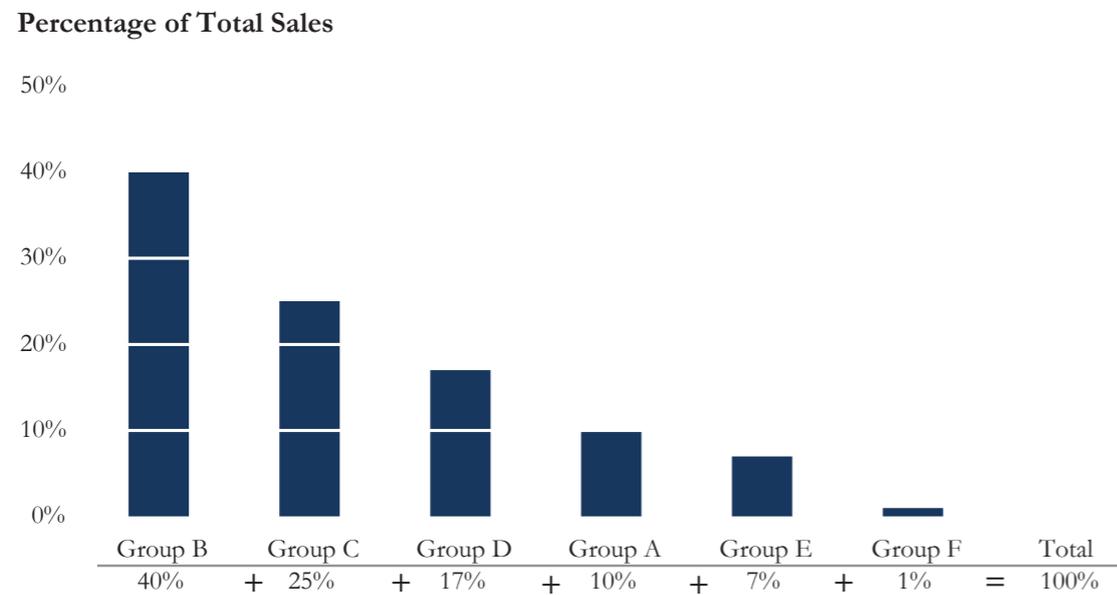


Figure 9C
Alternative to a Pie Chart: A Stacked Bar Chart

Shares of Aggregate Income, 1962 and 2009
(Percent)

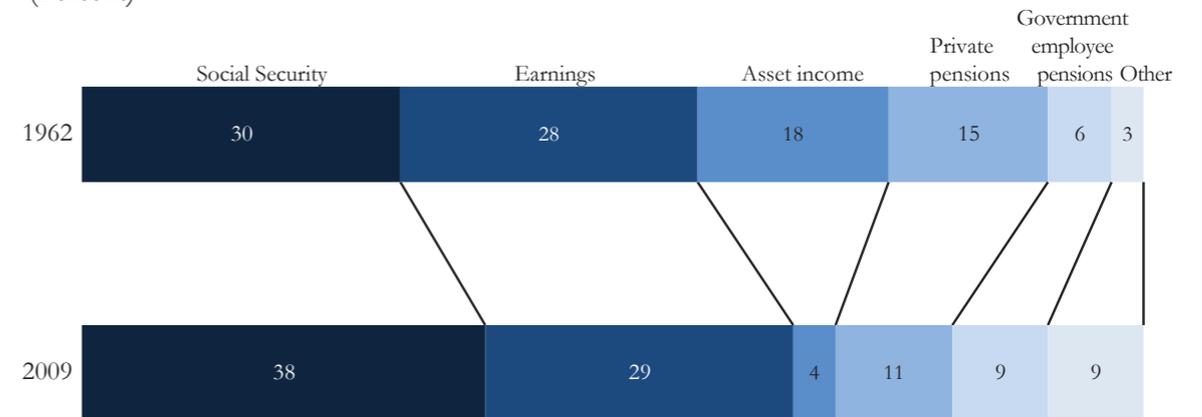
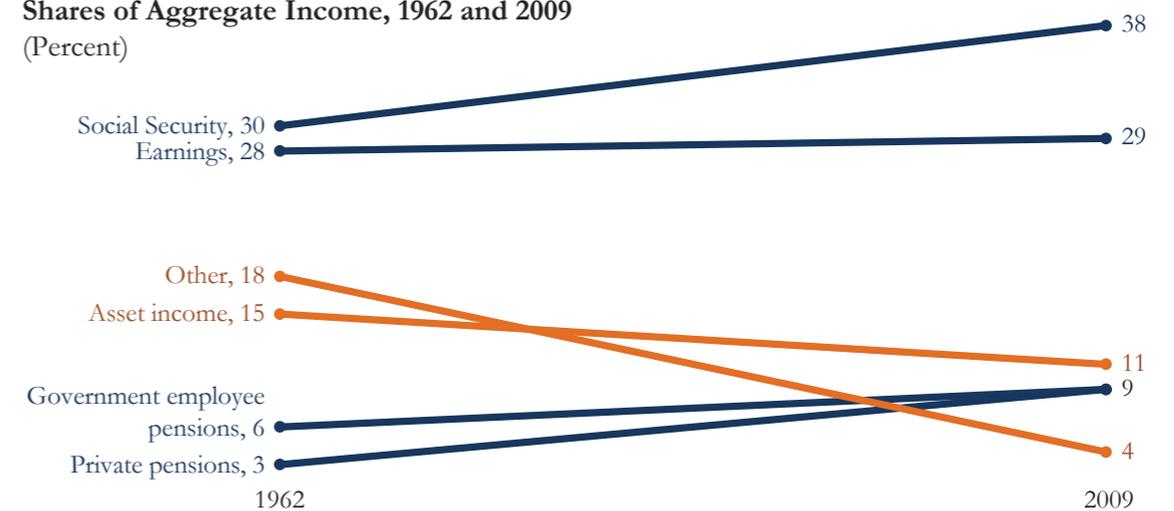


Figure 9D
Alternative to a Pie Chart: The Slope Chart

Shares of Aggregate Income, 1962 and 2009
(Percent)



Schwabish, J. (2014). An Economist's Guide to Visualizing Data Journal of Economic Perspectives 28(1), 209 - 234.

„how to make a manipulative chart“
Graphische Manipulation

Die Wahrnehmung der Ungleichheit

- Kritische Punkte:
 - Asymmetrische Behandlung der Achsen (siehe Spitze) – Effekt für Skalierung
 - Verwendung suggestiver Idealtypen ohne Achsen oder Orientierungspunkte (zB Median).
 - Graphische Übereinstimmung mit Idealtypen hängt ab von der Höhe der Balken.

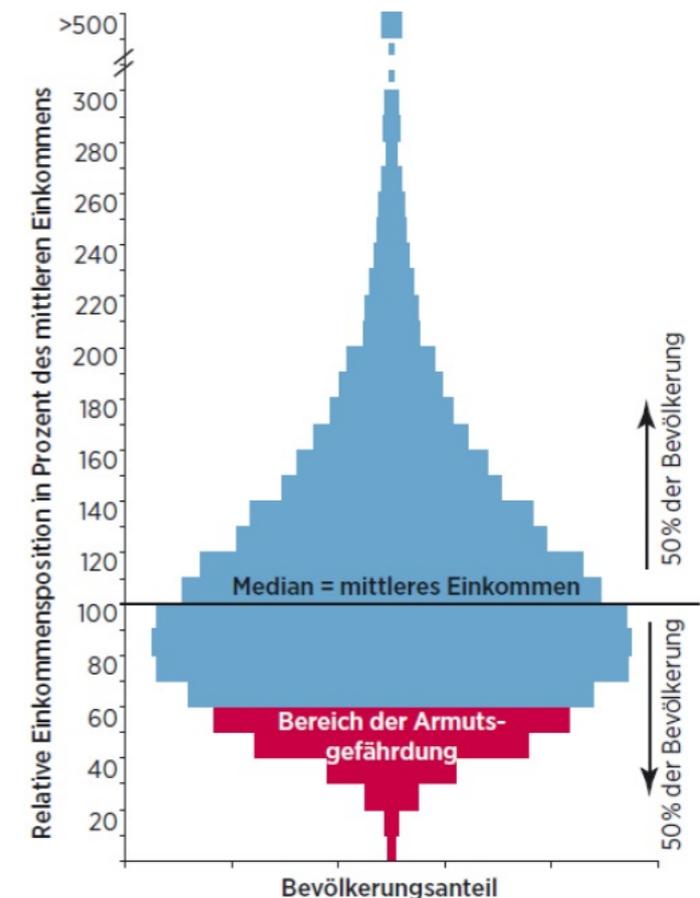
Ungleichheit: Wahrnehmung und Wirklichkeit in Deutschland

Wie die Deutschen die Gesellschaft sehen ...
Zustimmung im Jahr 2009 in Prozent

	Eine kleine Elite oben, nur sehr wenige Menschen in der Mitte und die große Masse der Bevölkerung unten.	18,8
	Eine Gesellschaft, die einer Pyramide gleicht, mit einer kleinen Elite oben, mehr Menschen in der Mitte und den meisten Menschen unten.	35,4
	Eine Gesellschaft, die einer Pyramide gleicht, aber mit nur wenigen Menschen ganz unten.	23,0
	Eine Gesellschaft, in der sich die meisten Menschen in der Mitte befinden.	18,6
	Viele Menschen im oberen Bereich und nur wenige Menschen im unteren Bereich.	4,2

Handelsblatt | Quellen:
ISSP, Eurostat (EU-SILC)

... und wie die Einkommensschichtung tatsächlich beschaffen ist
Bevölkerungsanteile im Jahr 2009 in Prozent



Lesebeispiel: Der rote Bereich der Armutsgefährdung gibt den Anteil der Bevölkerung an, der ein bedarfsgewichtetes Nettoeinkommen von weniger als 60 Prozent des mittleren Einkommens (Median) der Gesamtbevölkerung zur Verfügung hat (15,6 %).

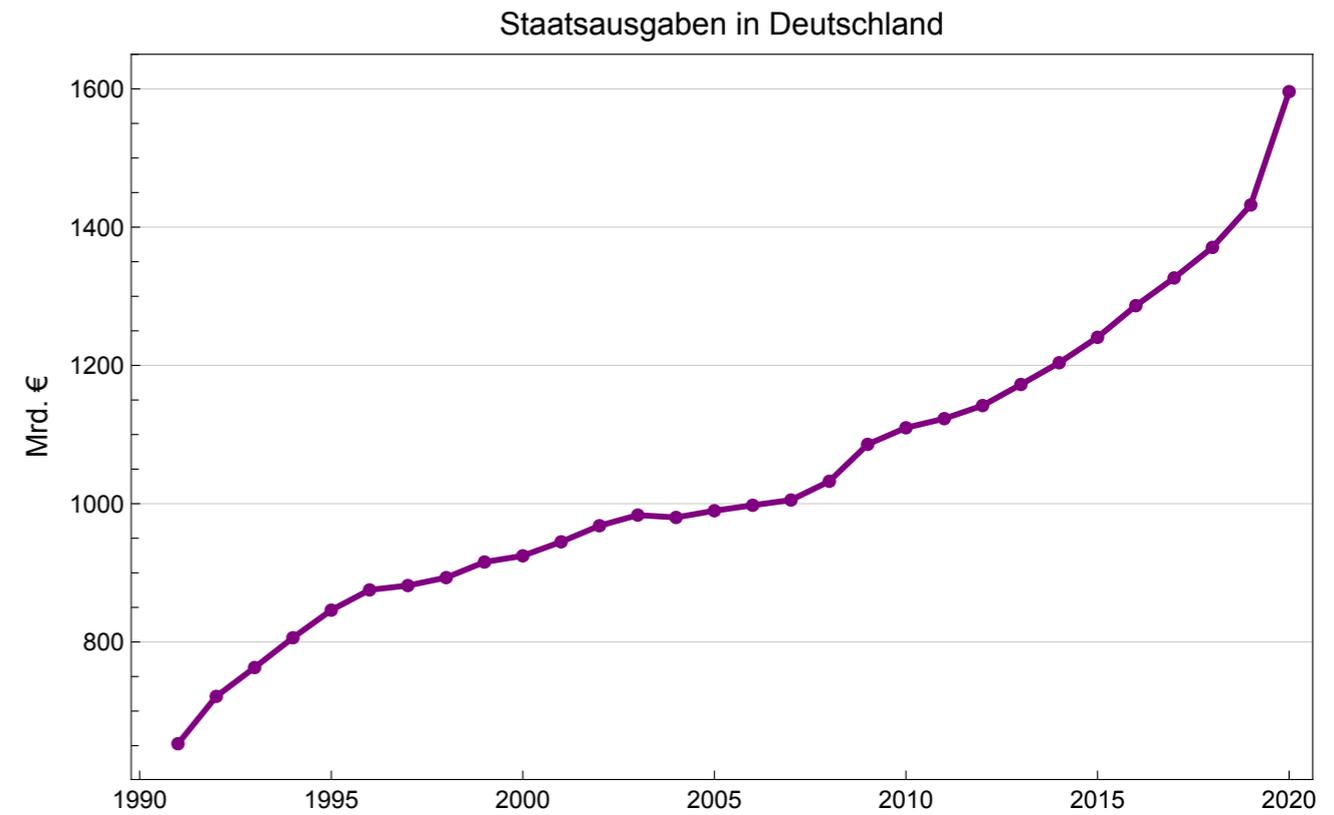
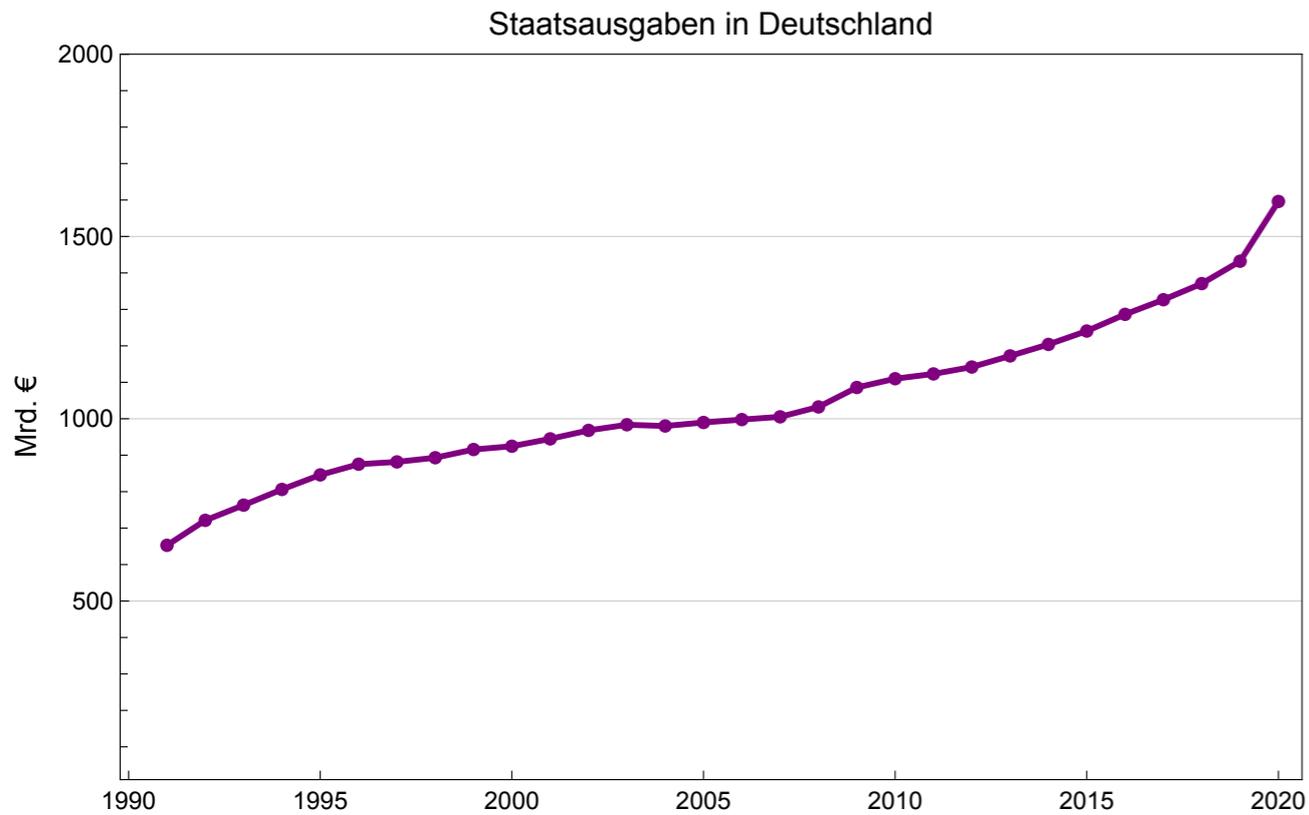
https://de.ecedon.uni-oldenburg.de/fyls/25189/download_file/

Grafische Manipulation: Beispiele

- Achsenvariationen:
 - Vertikale Achse mit möglichst hohem (bzw. niedrigem) Wert beginnen und mit möglichst niedrigem (bzw. hohem) Wert enden lassen.
 - Horizontale Achse stauchen.
 - Auslassungen von Achsenabschnitten oder uneinheitliche Achsen.
- Zeitliche Spielereien
 - Zeitreihen mit Aufwärtsbewegung (bzw. Abwärtsbewegung) aber ohne kausalen Zusammenhang auf der horizontalen und vertikalen Achse abbilden, um einen Kausalzusammenhang zu suggerieren.
 - Ähnlich verlaufende Zeitreihen als quasi-kausal nebeneinanderstellen.
- Irreführende Symbole, räumliche Darstellungen

Beispiel: Stauchung der vertikalen Achse

- Wie dramatisch ist der Anstieg der Staatsschulden?

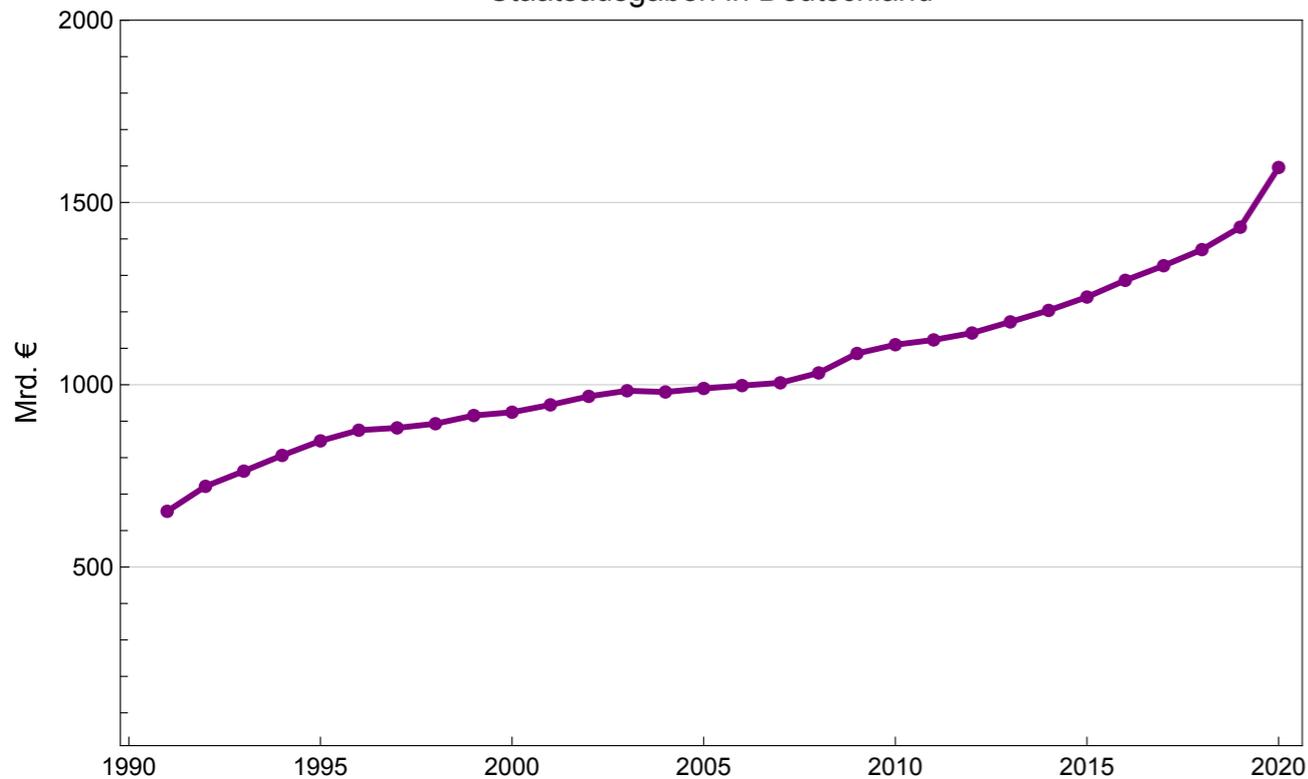


Quelle: AMECO, Total current expenditure: general government :- Excessive deficit procedure

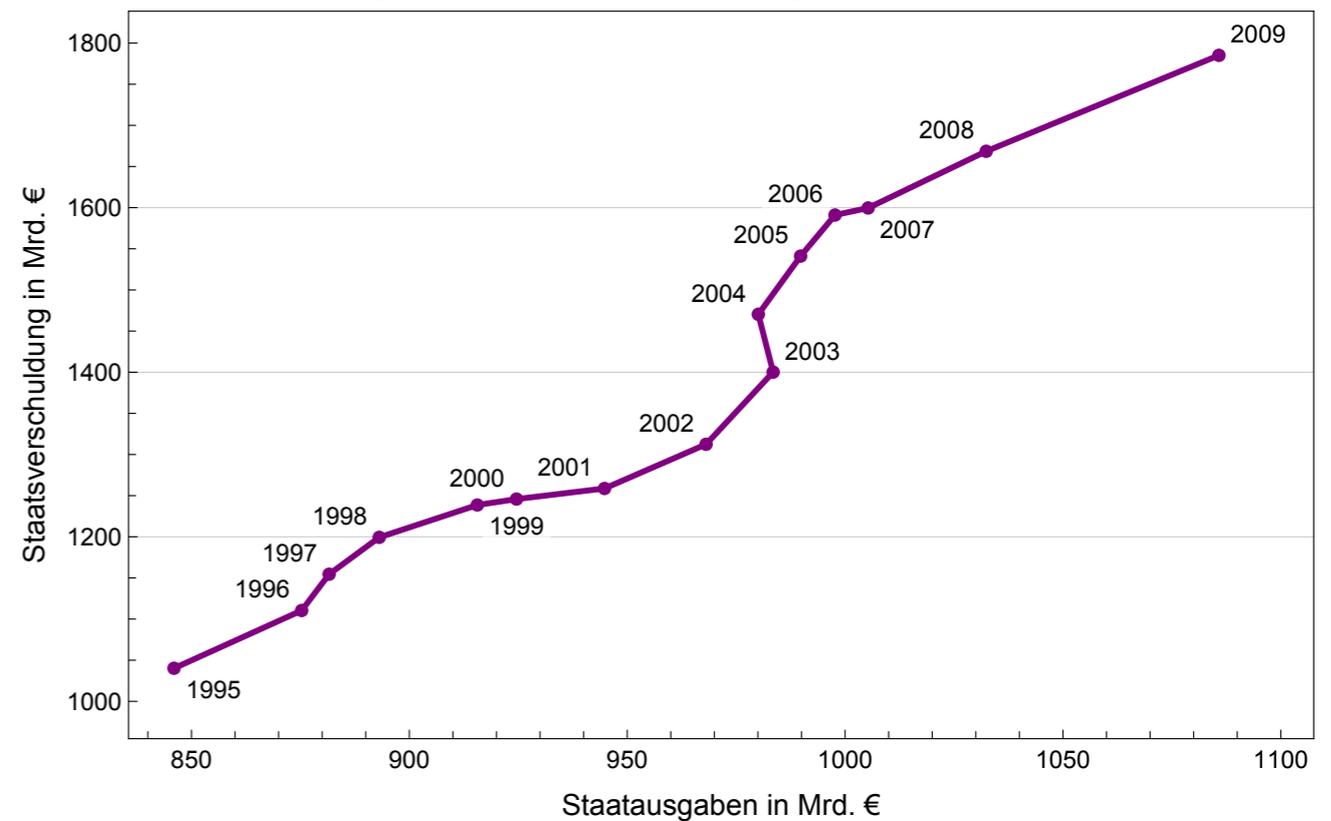
Beispiel: Suggestives Kombinieren von Variablen

- Treiben die Staatsausgaben die Staatsschulden?

Staatsausgaben in Deutschland



Staatsausgaben und Staatsverschuldung in Deutschland (1995–2009)

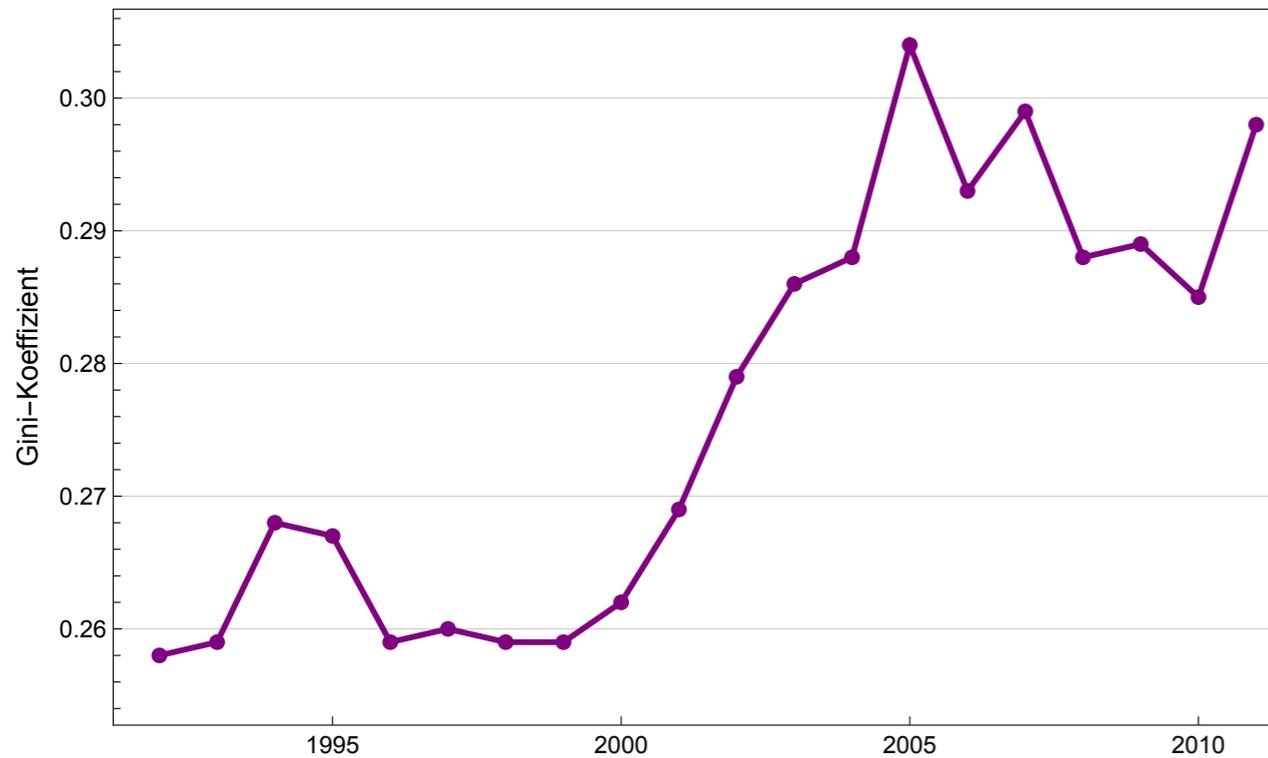


Quelle: AMECO, Total current expenditure: general government :- Excessive deficit procedure

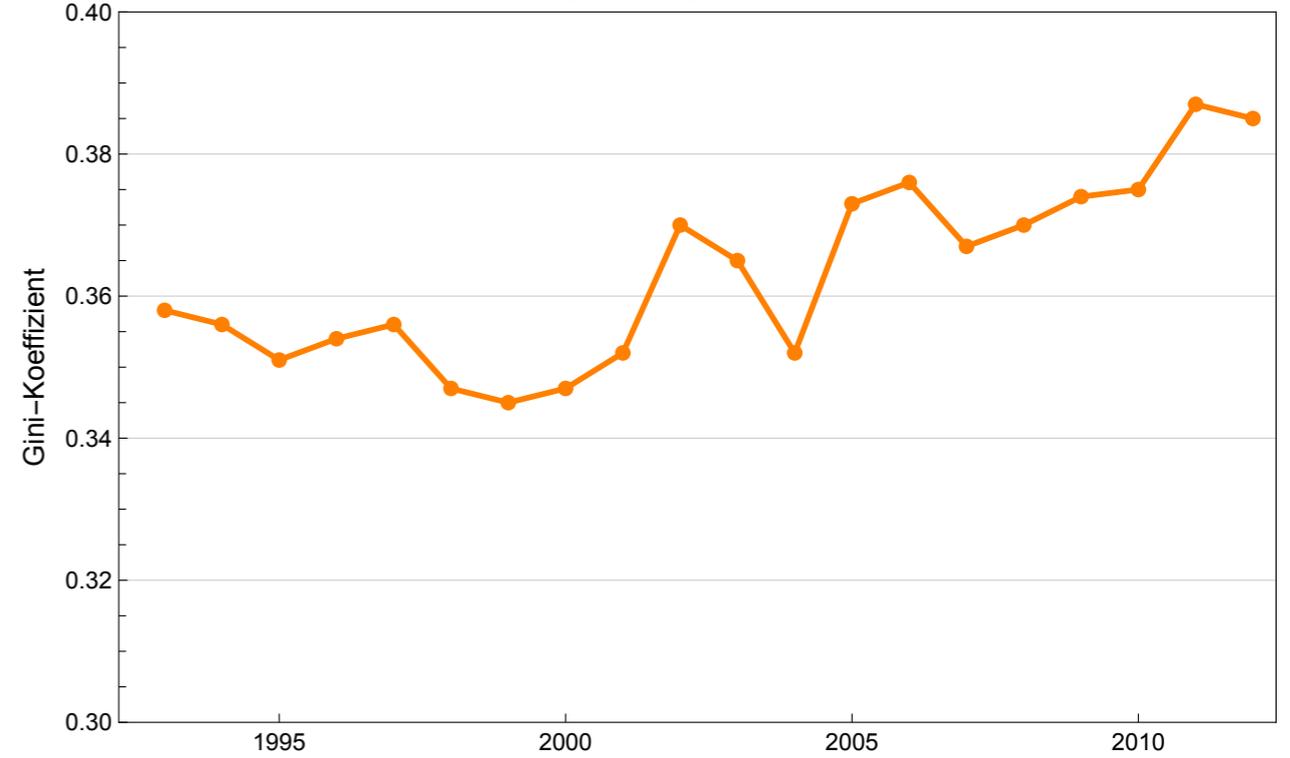
Deutschland als „Mutterland der Ungleichheit“

Ein manipulativer Vergleich mit den USA

Entwicklung der Ungleichheit in Deutschland



Entwicklung der Ungleichheit in den USA

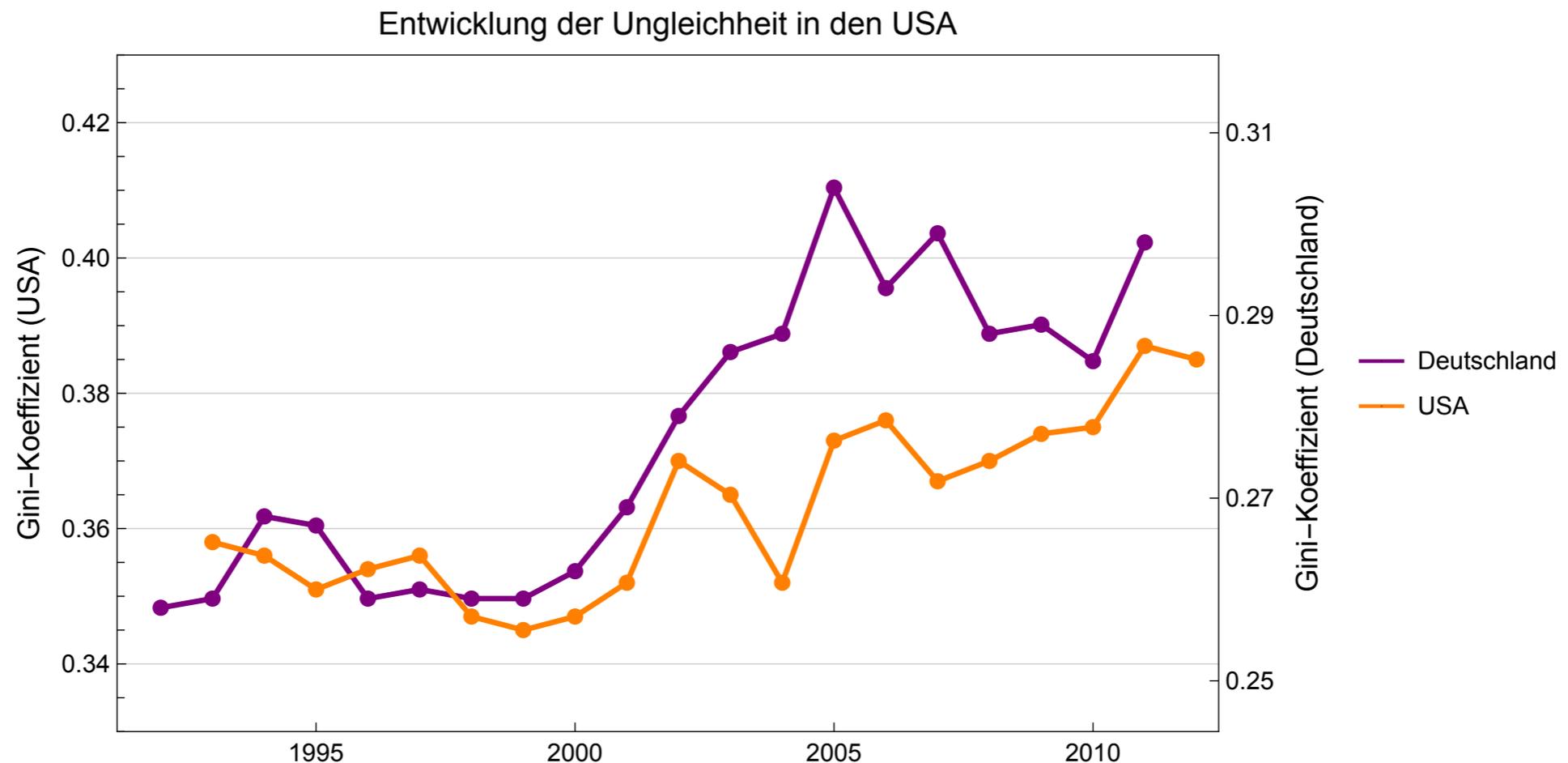


Gini-Koeffizient, äquivalenzgewichtete verfügbare Haushaltseinkommen, 1992-2012
Quelle: OECD.

Deutschland als „Mutterland der Ungleichheit“

Ein (anderer) manipulativer Vergleich mit den USA

- Ist die Ungleichheit in Deutschland wirklich höher?
 - Man beachte die zweite Achse auf der rechten Seite!

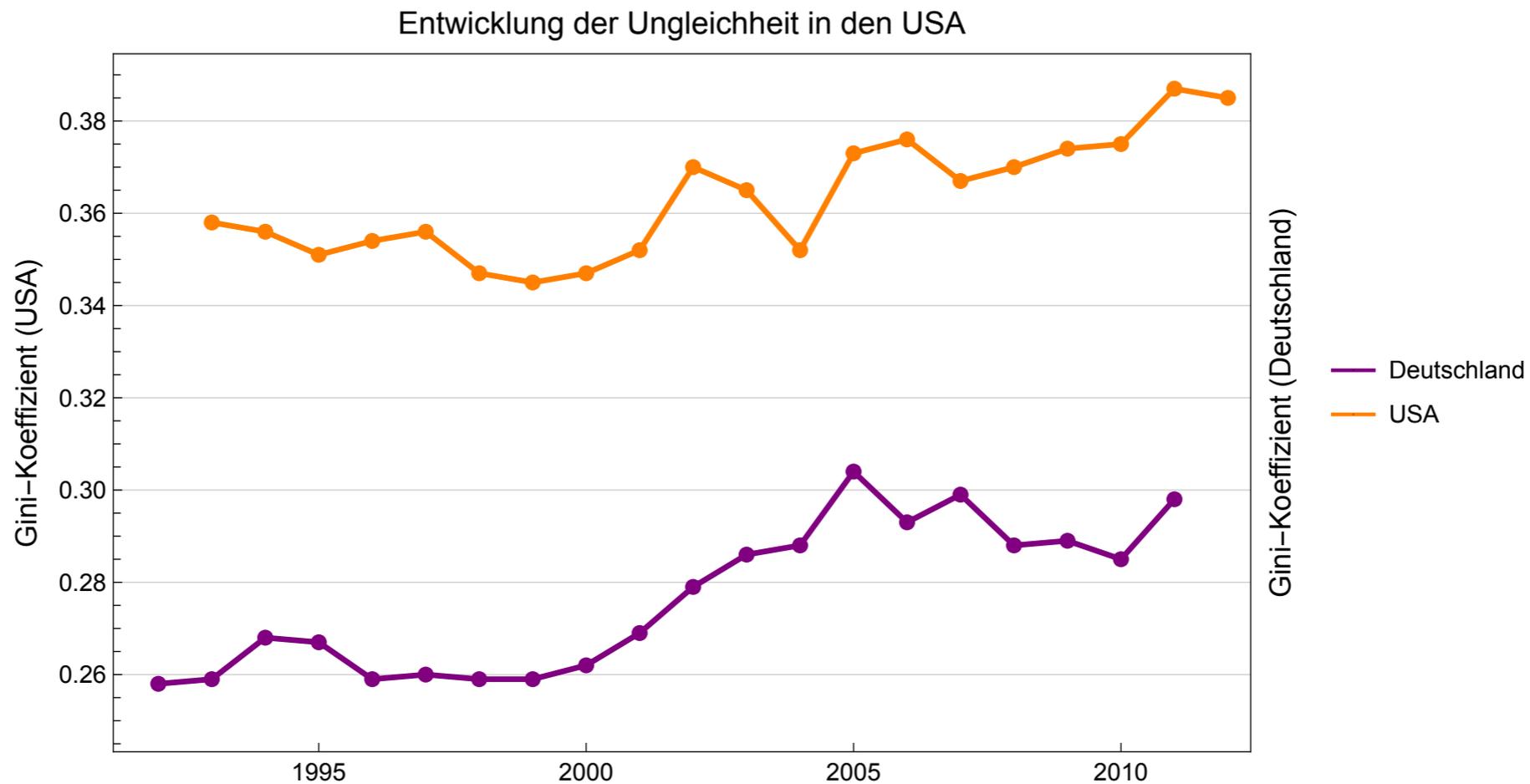


Gini-Koeffizient, äquivalenzgewichtete verfügbare Haushaltseinkommen, 1992-2012
Quelle: OECD.

Deutschland als „Mutterland der Ungleichheit“

Gleichmäßig skalierte Graphen zeigen folgendes Bild

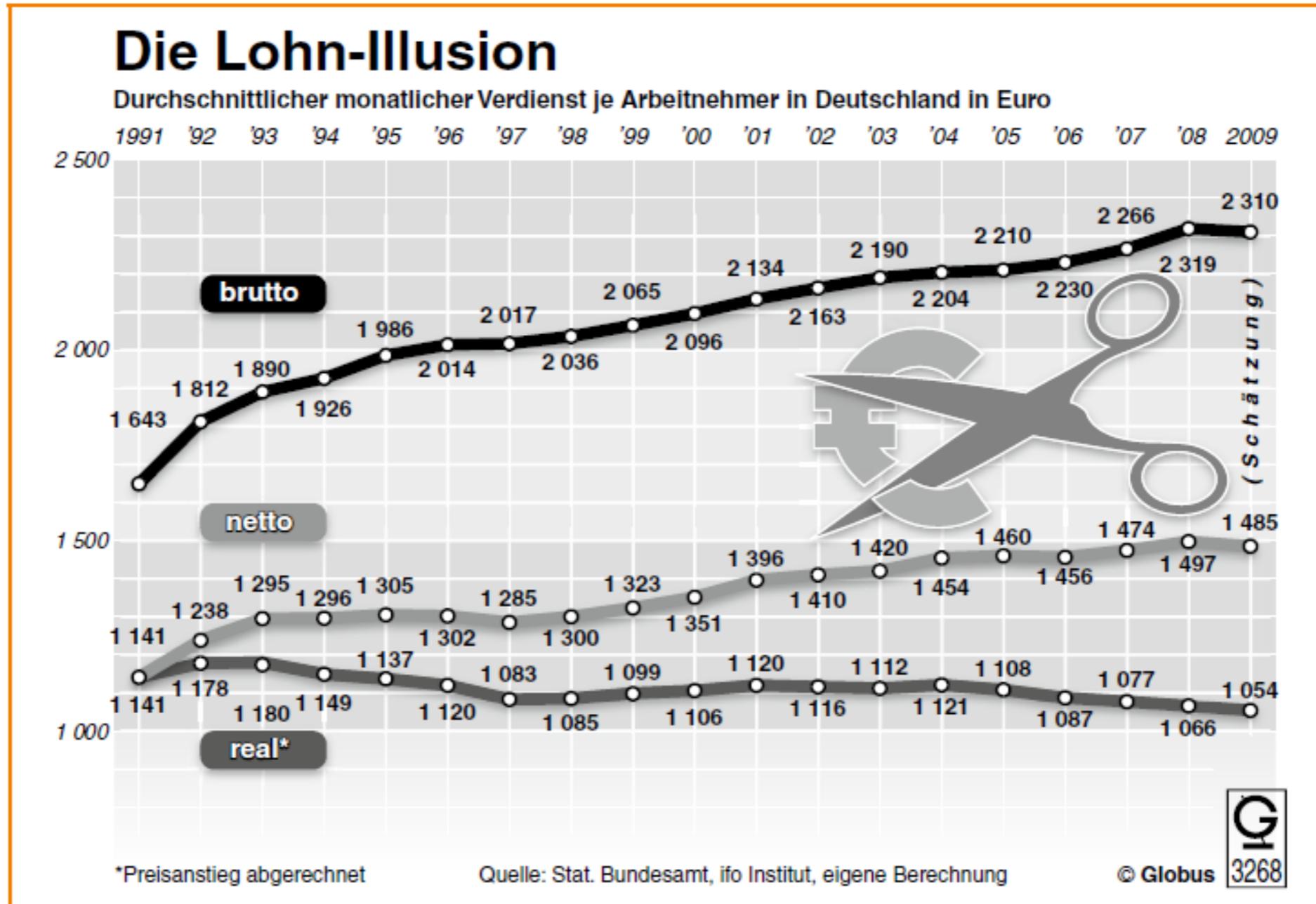
- Absoluter Anstieg des Gini (Minimum zu Maximum) ungefähr gleich (+0.04).
 - Ausgangsniveau in D wesentlich niedriger → stärkerer *relativer* Anstieg.



Gini-Koeffizient, äquivalenzgewichtete verfügbare Haushaltseinkommen, 1983-2010
Quelle: OECD.

Der Staat „beschneidet“ unser Einkommen?

Einkommensentwicklung



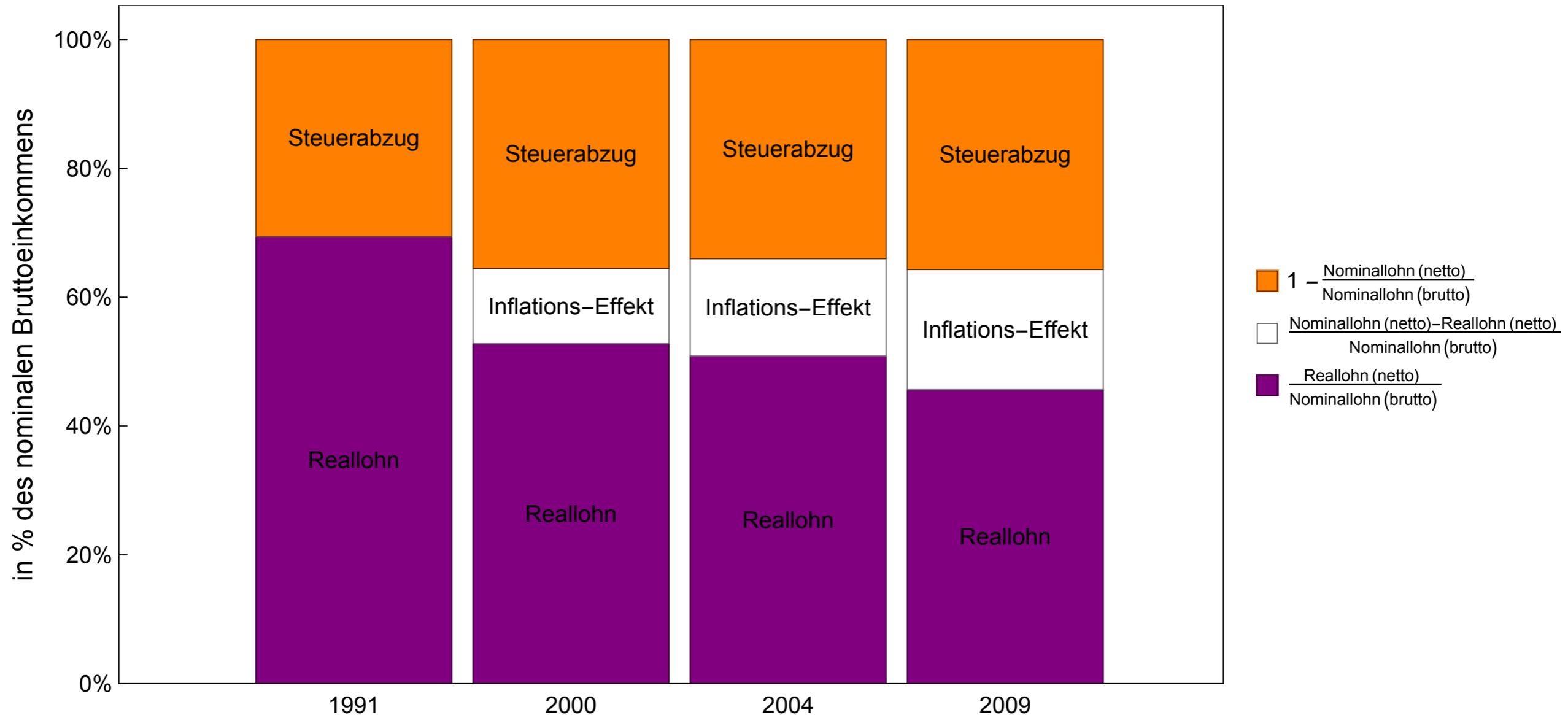
Nominale Brutto- und Nettolöhne, Reallöhne

Quelle: http://www.handelsblattmachtschule.de/fileadmin/PDF/UE_Finanz-Allgemein_2011_final-online.pdf

Was sagen eigentlich die Zahlen dazu?

Passen Werte und Symbolik zusammen?

Woher kommt die Schere?



Quelle: Eigene Berechnung auf Basis von Handelsblatt macht Schule (2011)

„how to make a manipulative chart“
Statistische Manipulation

Statistische Manipulation

- „**Statistische Manipulation**“: Wahl einer statistischen Kennzahl und Ausdruckweise, die den zu Grunde liegenden Sachverhalt dramatisiert.
 - Statistische Manipulation als „pointierte Zuspitzung“ \neq glatte Lüge
- Beispiele:
 - Entwicklung einer Quote (die bereits in % angegeben wird) als prozentuale Veränderung zeigen (d.h. Erstellung eines Index, der im Basisjahr auf 100 normiert wird).
 - Für die Zukunft prognostizierte Entwicklungen mit in die Darstellung aufnehmen, wenn sie die Abbildung in gewünschter Weise dramatischer aussehen lassen.
 - Gewisse Blickwinkel auf die Daten betonen, andere vernachlässigen, z.B. Verhältnis von absoluten und relativen Unterschieden.

Wie ließe sich der folgende Sachverhalt dramatisieren?

	Lohnquote
2005	60%
2015	54%

Die Lohnquote ist der Anteil der Lohneinkommen am gesamten Volkseinkommen.

„Die Lohnquote ist um 10% gefallen!“

	Lohnquote
2005	60%
2015	54%

Die Lohnquote ist der Anteil der Lohneinkommen am gesamten Volkseinkommen.

Wie könnte eine weniger dramatische Aussage lauten?

	Lohnquote
2005	60%
2015	54%

Die Lohnquote ist der Anteil der Lohneinkommen am gesamten Volkseinkommen.

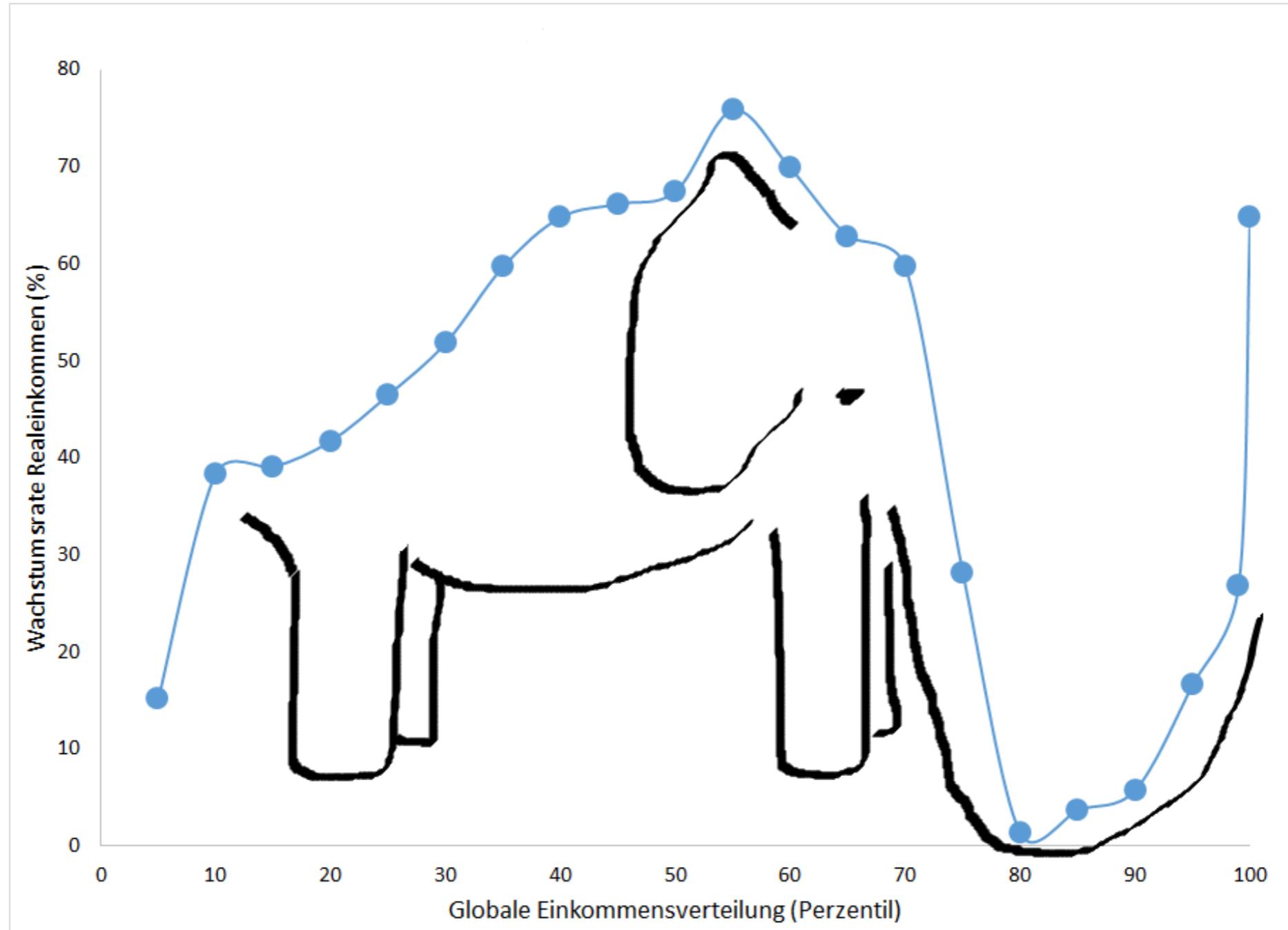
Die Lohnquote ist um 6 Prozentpunkte gefallen.

	Lohnquote
2005	60%
2015	54%

Die Lohnquote ist der Anteil der Lohneinkommen am gesamten Volkseinkommen.

Globalisierung und Einkommensverteilung

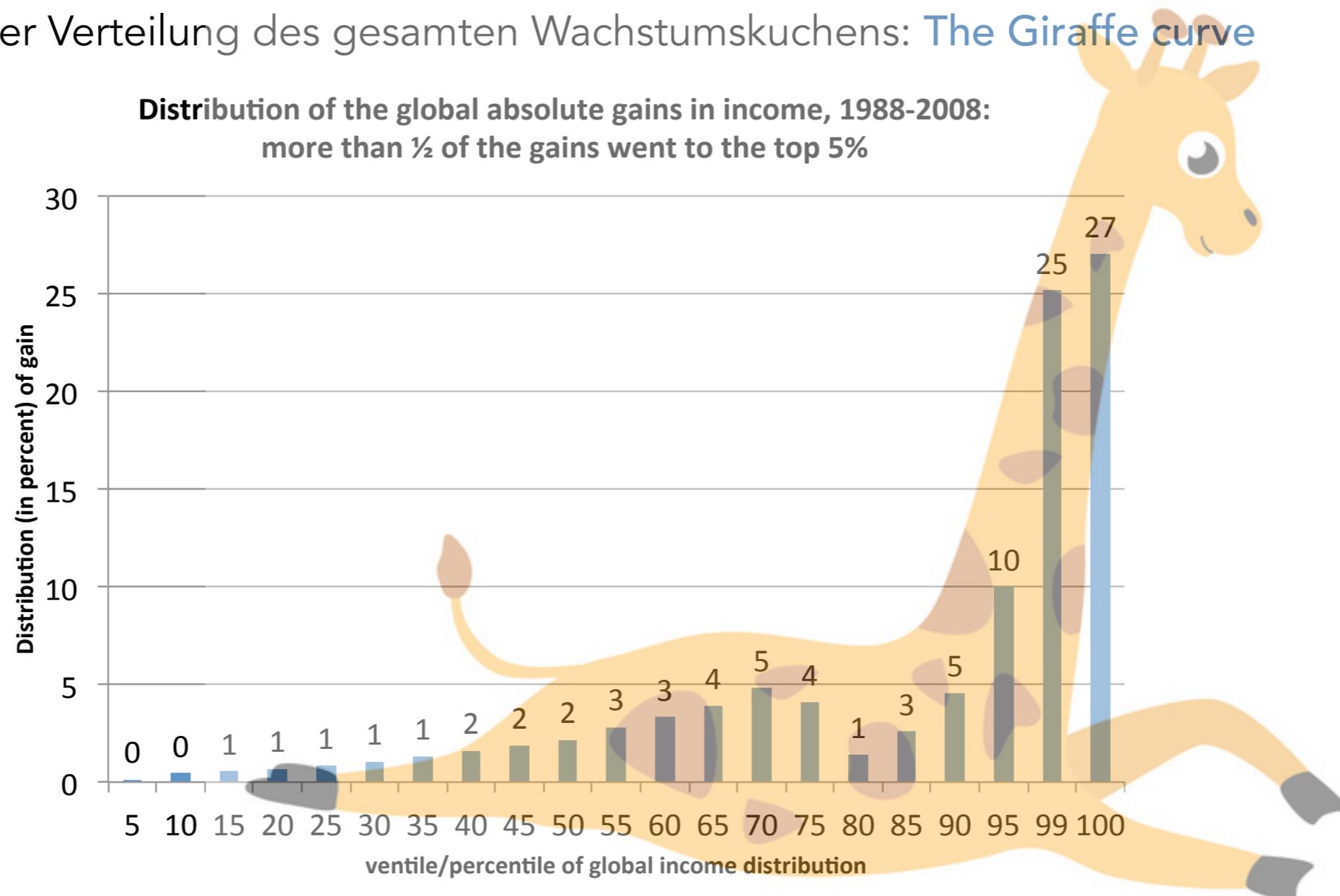
Relativer Einkommenszuwachs: „The Elephant Curve“



Globalisierung und Einkommensverteilung

Verteilung des absoluten Einkommenszuwachs: „The Giraffe Curve“

- Globale Einkommensverteilung zwischen 1988 und 2008.
 - Analyse der Verteilung des gesamten Wachstumskuchens: [The Giraffe curve](#)



Milanovic (2014), Lecture at the European Forum Alpbach

„how to make a manipulative chart“
Inhaltliche Manipulation

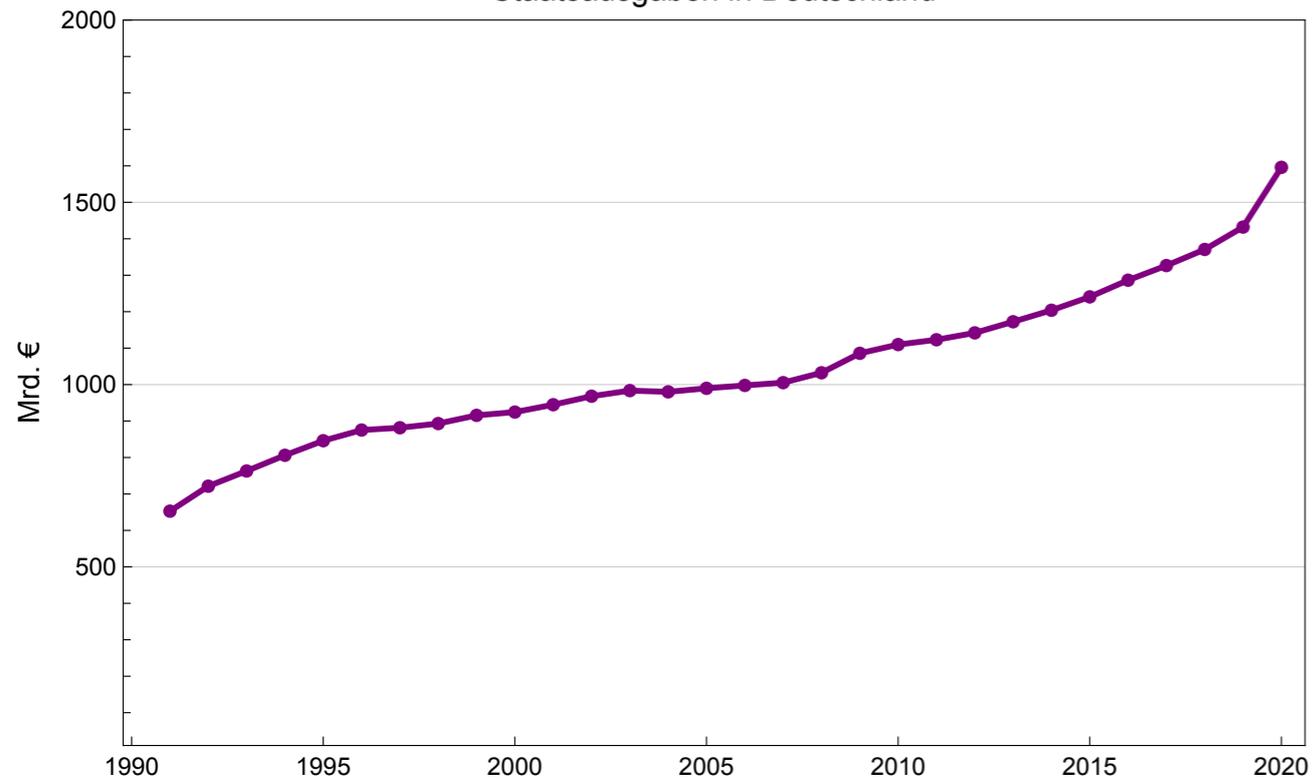
Inhaltliche Manipulation

- Anfangs- und Endzeitpunkt „geschickt“ auswählen, vorangehende und nachfolgende Entwicklung verbergen
 - Statt der gesamten Zeitreihe nur die Werte einiger weniger ausgewählter Zeitpunkte zeigen, die einen besonders steilen Anstieg oder Rückgang suggerieren und dazwischen liegende, weniger „passende“ Entwicklungen verschleiern.
- Nominale Werte statt reale Werte zeigen.
- Internationalen Vergleich je nach „Bedarf“ anpassen
- Irreführendes Nebeneinanderstellen von Abbildungen
- Von zwei oder mehreren Größen, die miteinander zusammenhängen, nur eine darstellen.

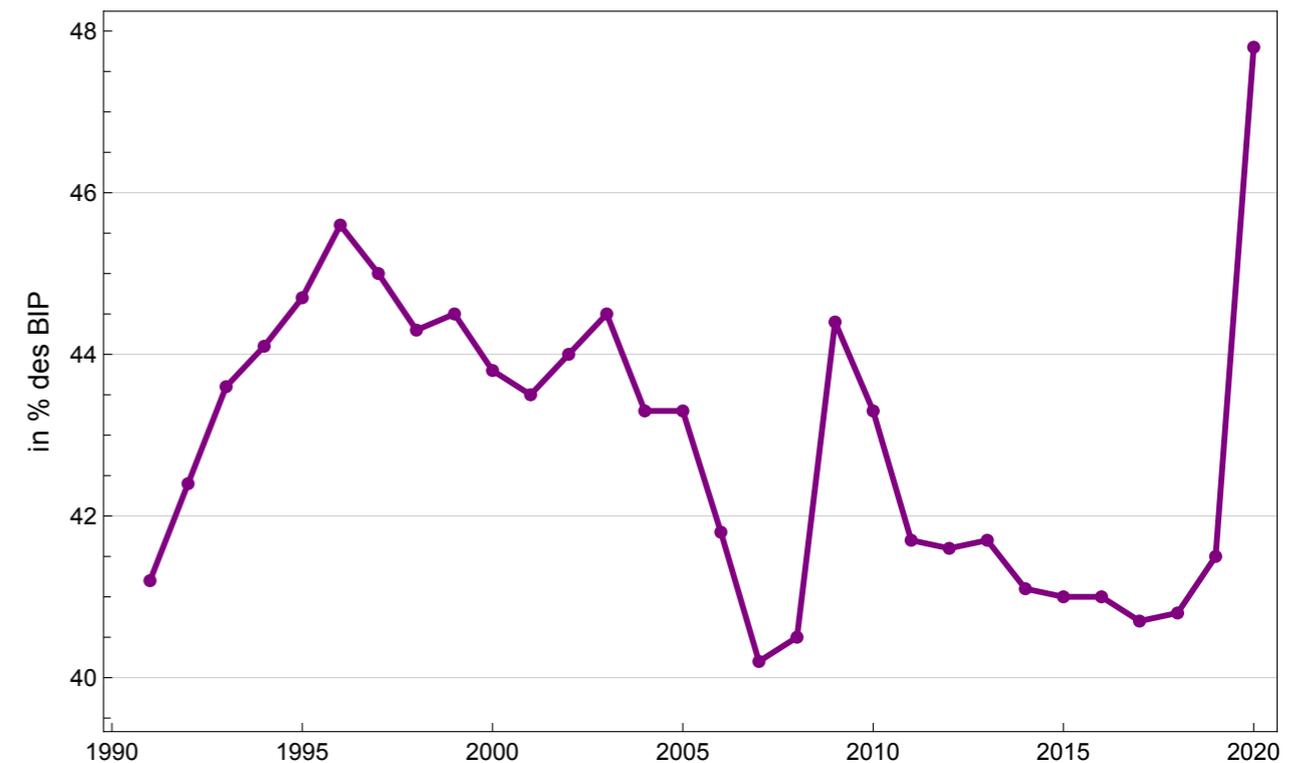
Beispiel: Weglassen zentraler Bezugsgrößen

- Schulden werden typischerweise in Relation zum laufenden Einkommen gesetzt.

Staatsausgaben in Deutschland



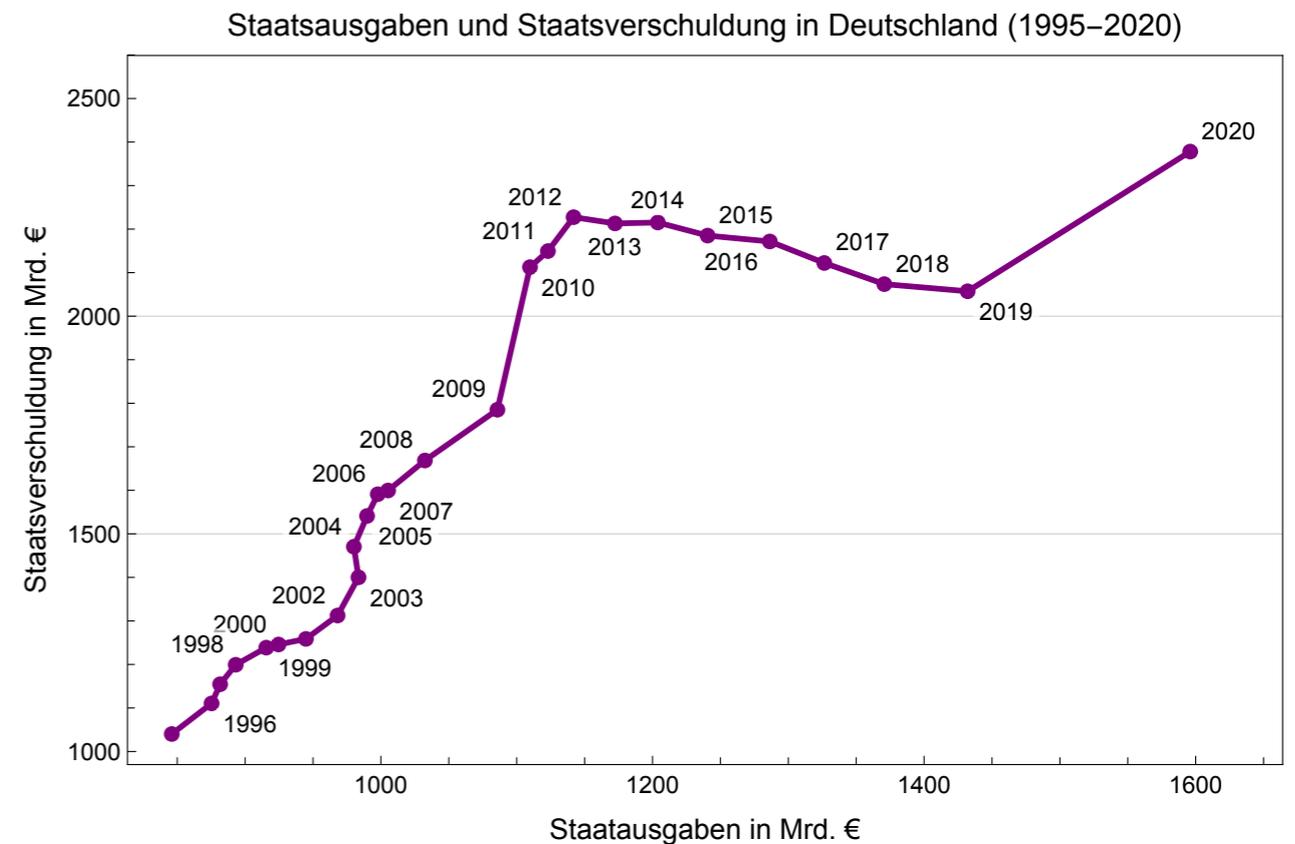
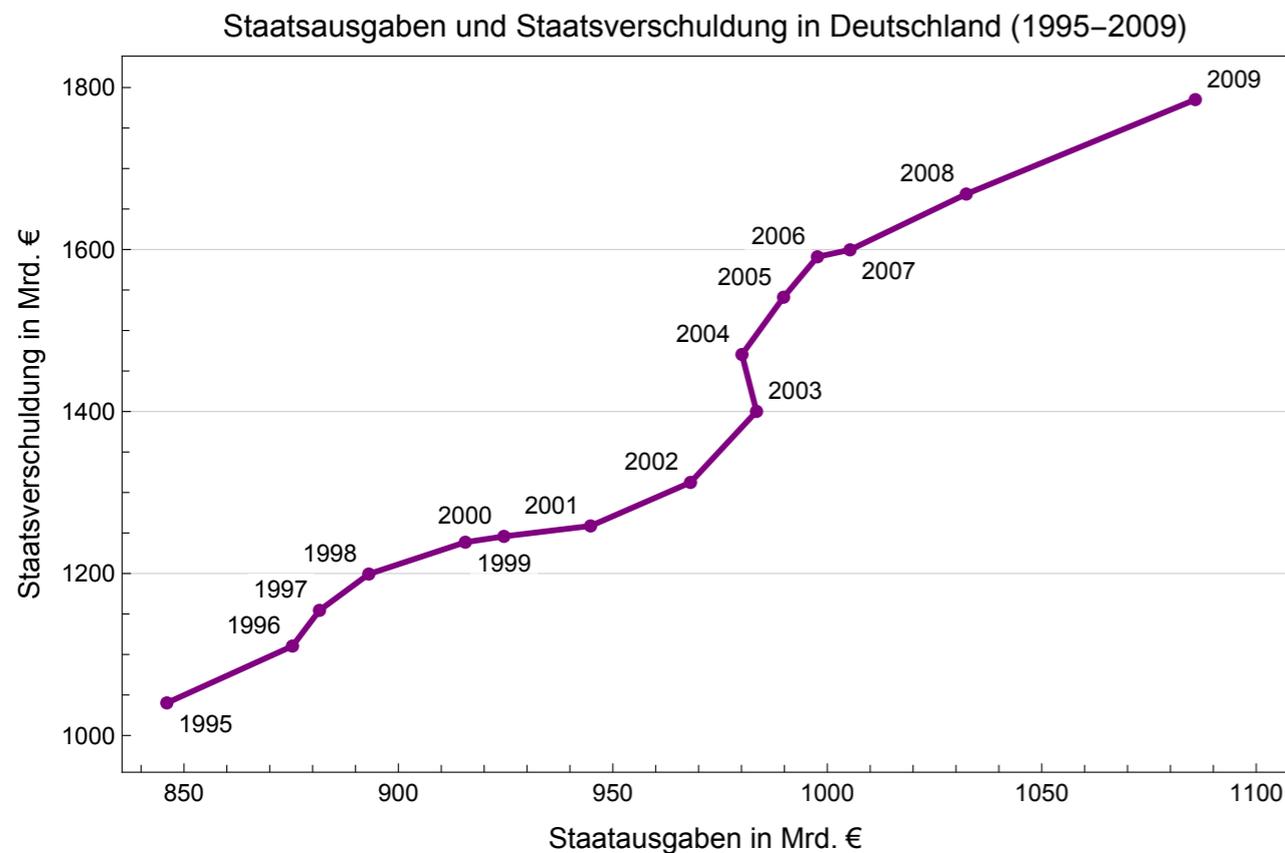
Staatsausgaben in Deutschland (relativ zum BIP)



Quelle: AMECO, Total current expenditure: general government :- Excessive deficit procedure

Beispiel: Suggestive Wahl des Beobachtungszeitraums

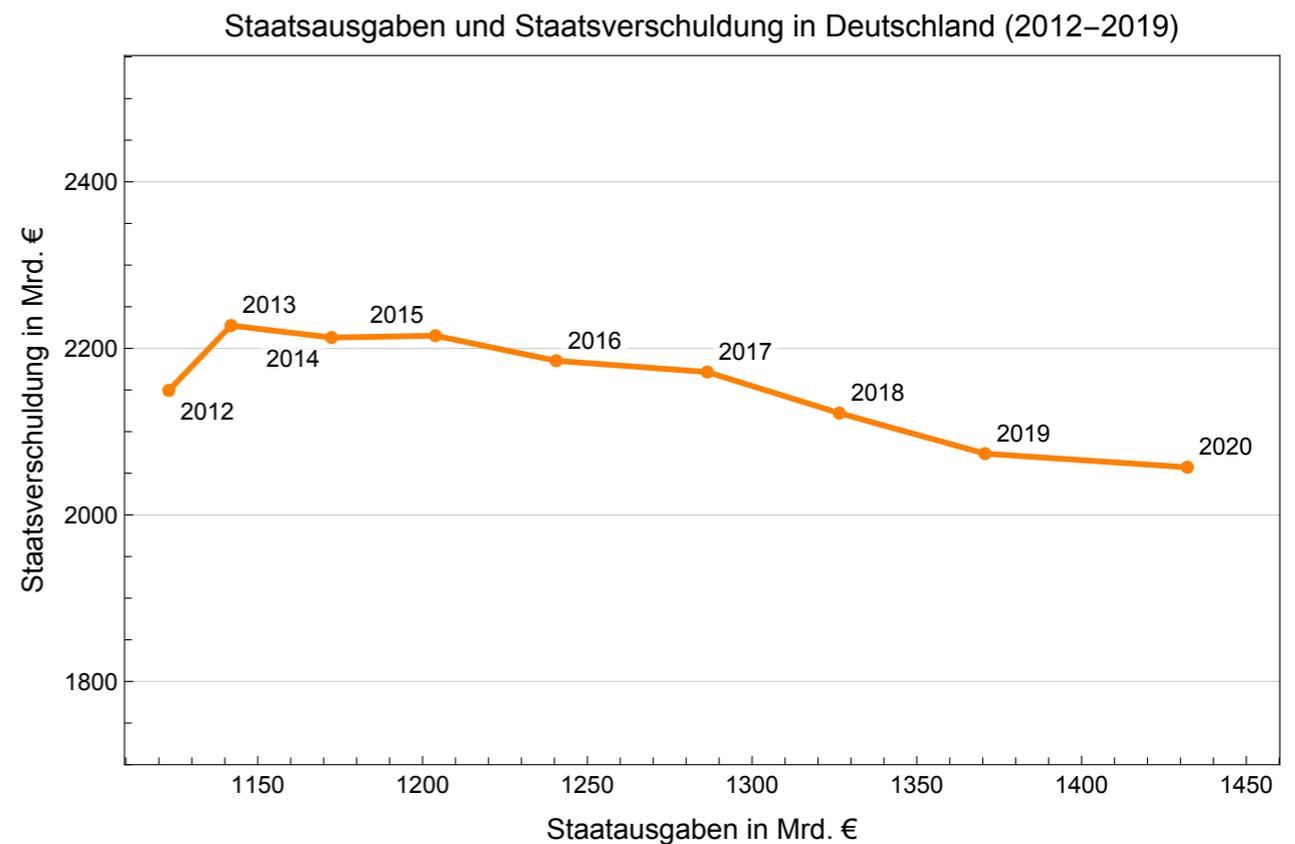
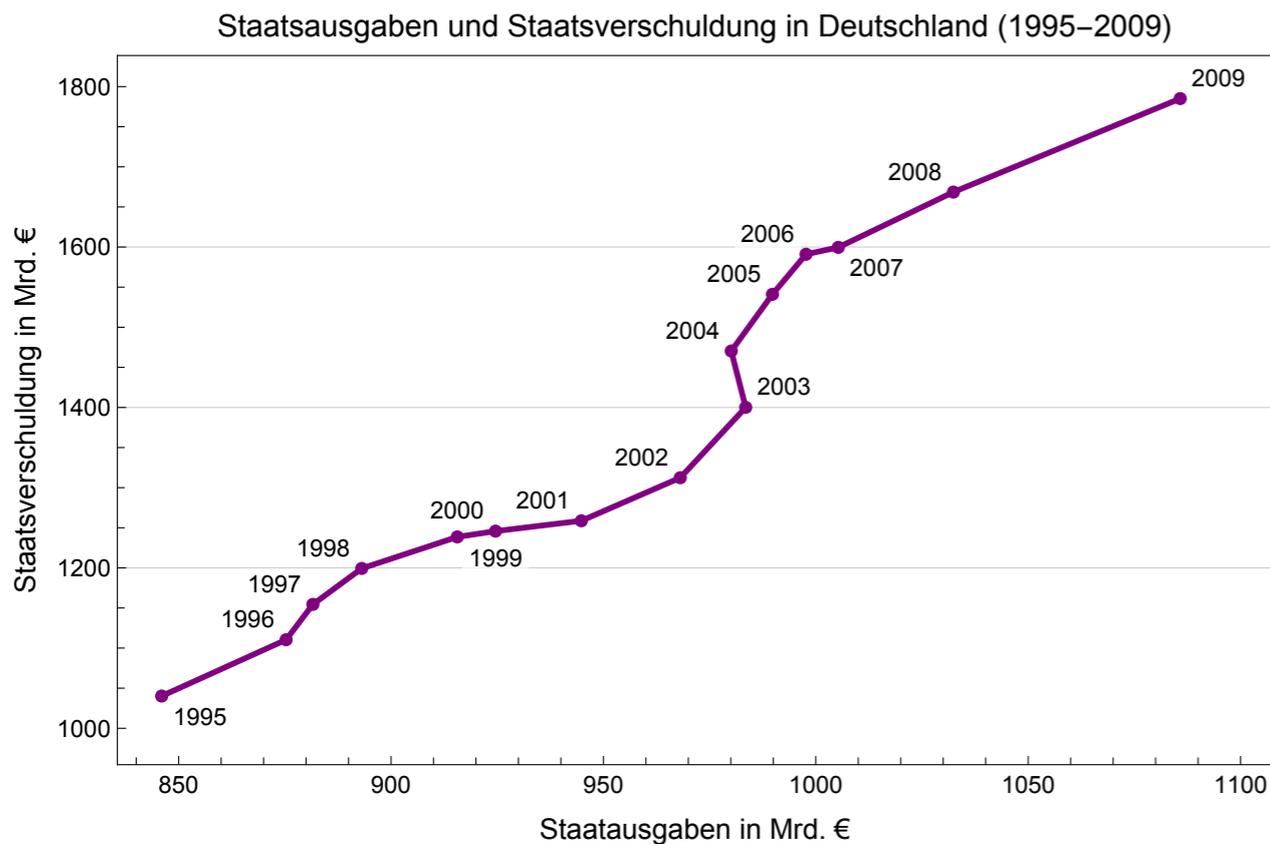
- Treiben die Staatsausgaben die Staatsschulden?



Quelle: AMECO, Total current expenditure: general government :- Excessive deficit procedure

Beispiel: Suggestive Wahl des Beobachtungszeitraums

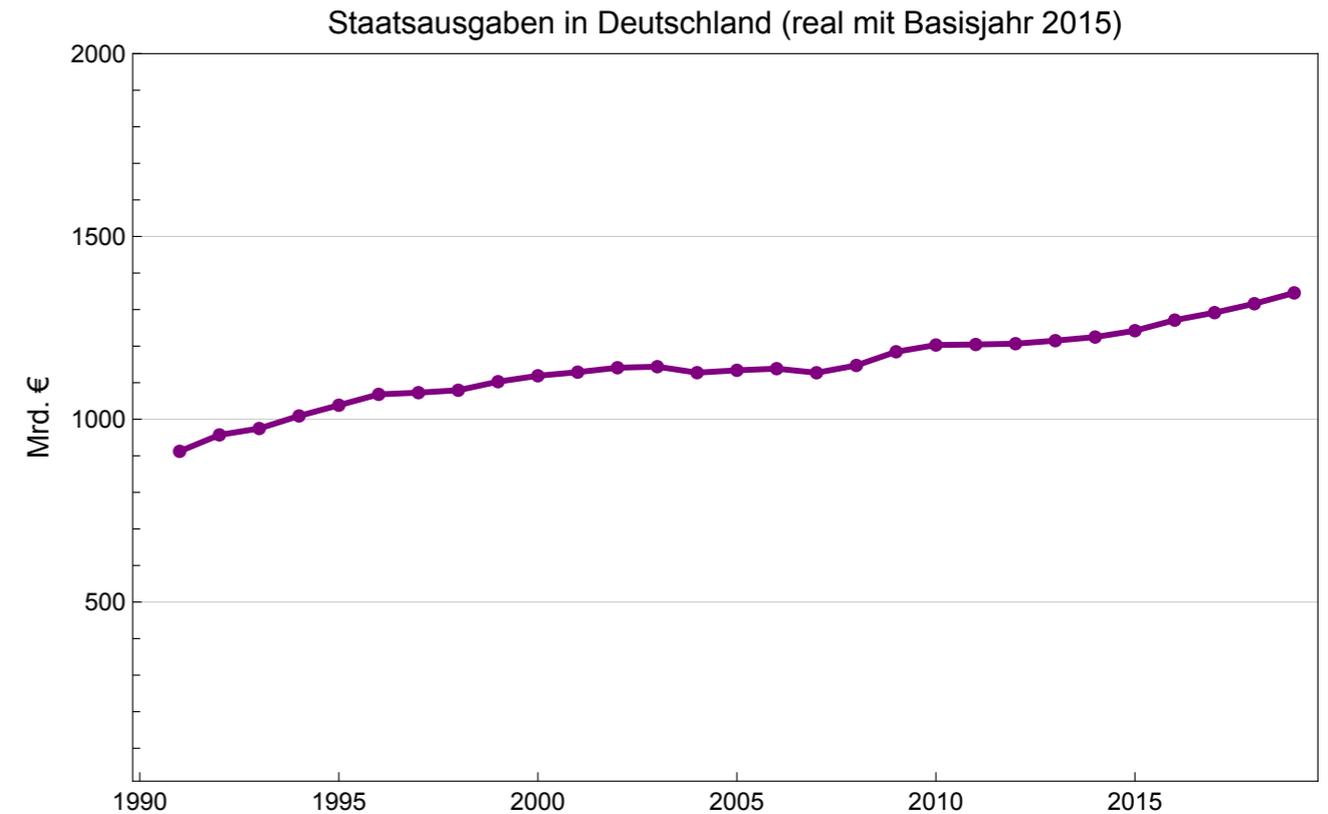
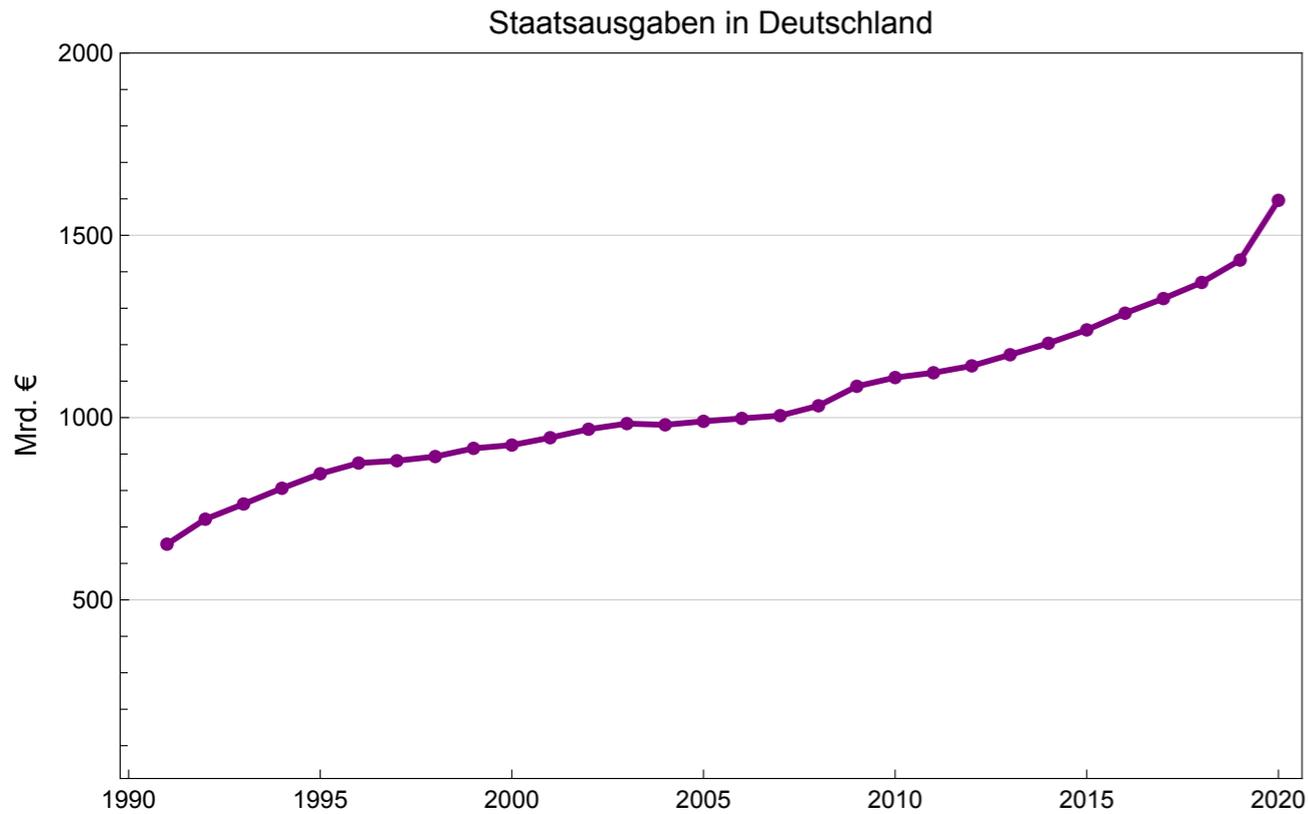
- Das zweite deutsche Wirtschaftswunder: Staatsausgaben senken Staatsschulden!



Quelle: AMECO, Total current expenditure: general government :- Excessive deficit procedure

Beispiel: Weglassen zentraler Einflussgrößen

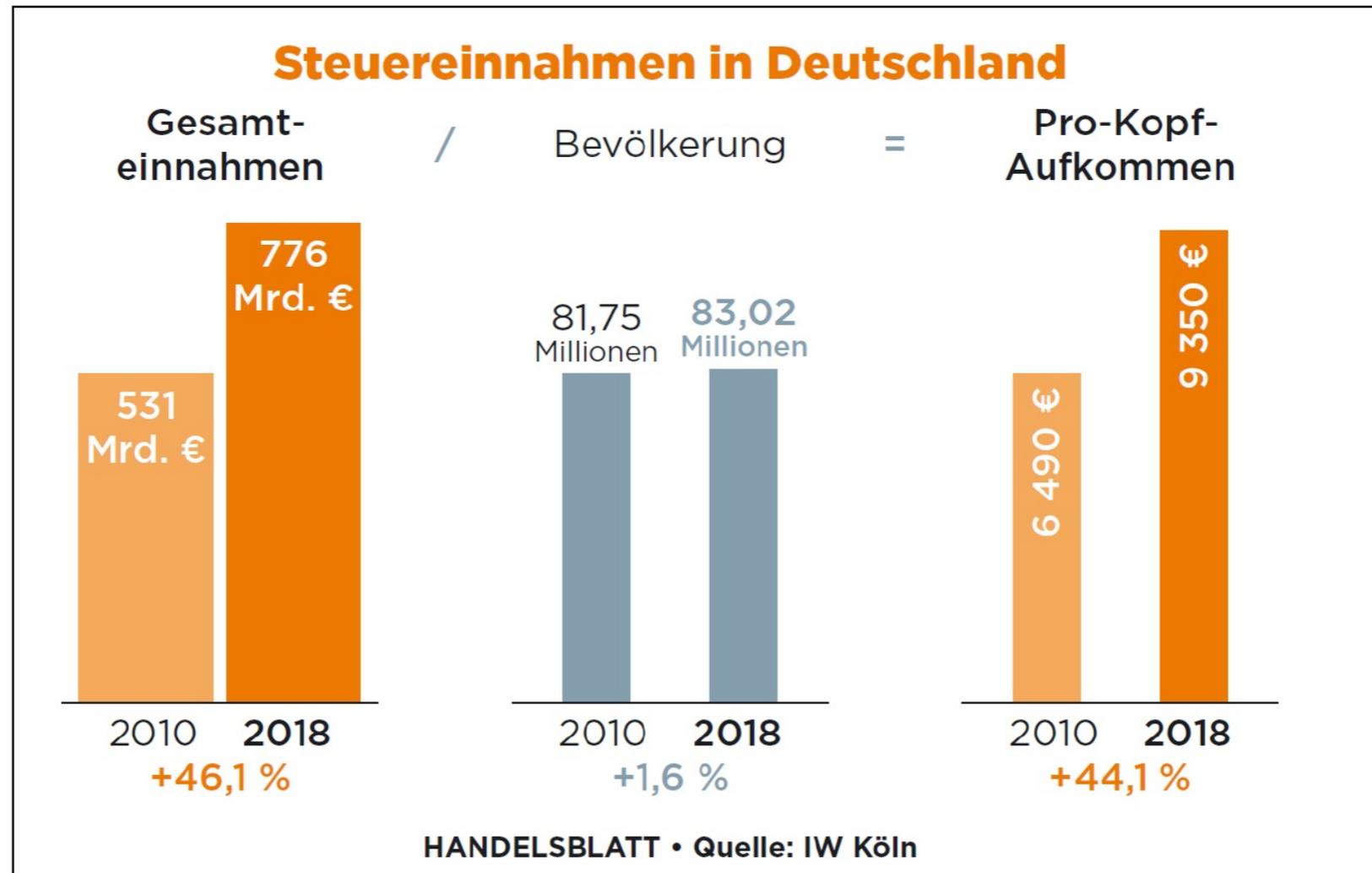
- Nominelle oder reale Staatsausgaben?



Quelle: AMECO, Total current expenditure: general government :- Excessive deficit procedure

Beispiel: Weglassen zentraler Einflussgrößen

- Gegenüberstellung mit Bevölkerungswachstum ist suggestiv
 - Sowohl Inflation als auch reales Wachstum werden ignoriert.

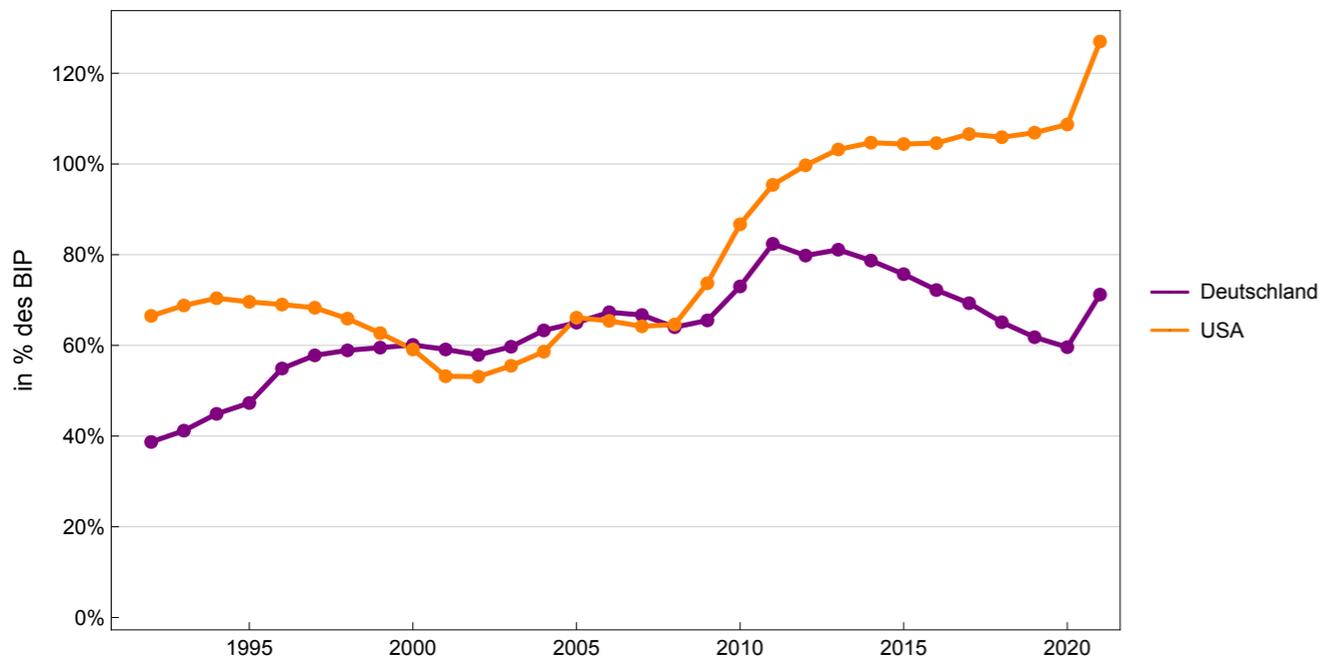


https://de.ecedon.uni-oldenburg.de/fyls/46712/download_file/

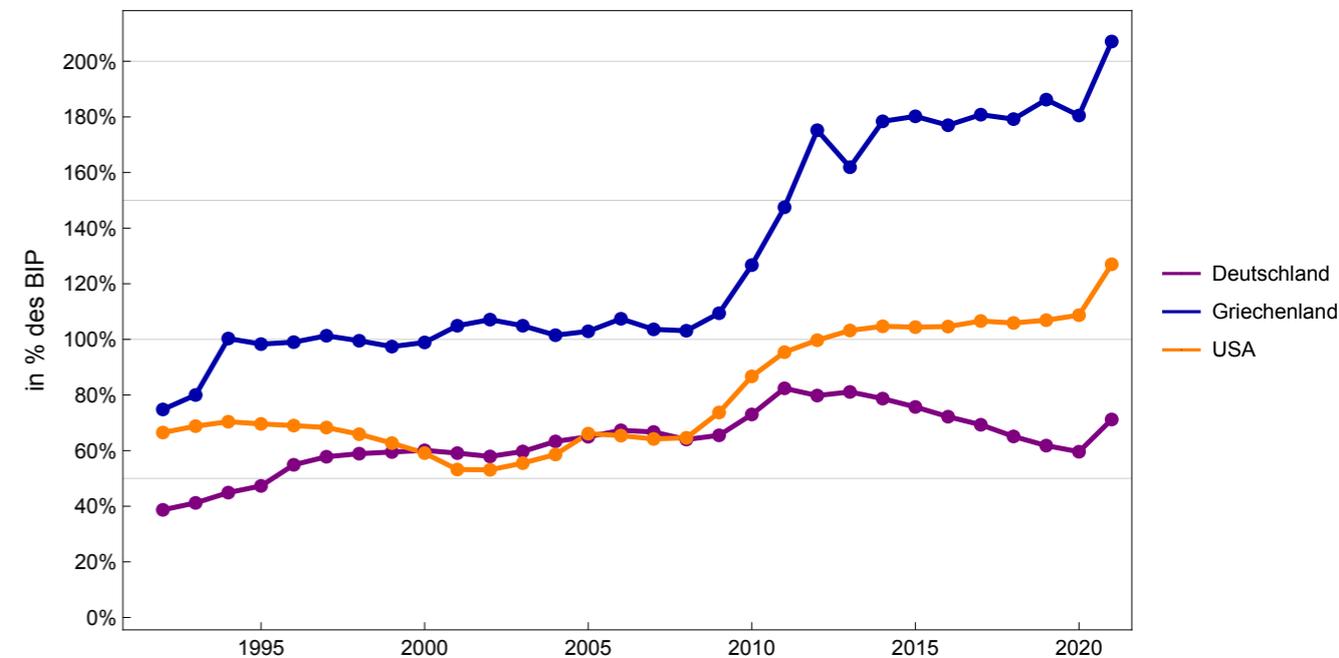
Gesucht: Spitzenreiter der Staatsverschuldung

- Internationale Vergleiche und die Wahl „passender“ Vergleichskandidaten
 - Diese prägen nicht uU nicht nur Minima und Maxima, sondern gesamte Skalierung.

Staatsverschuldung im internationalen Vergleich



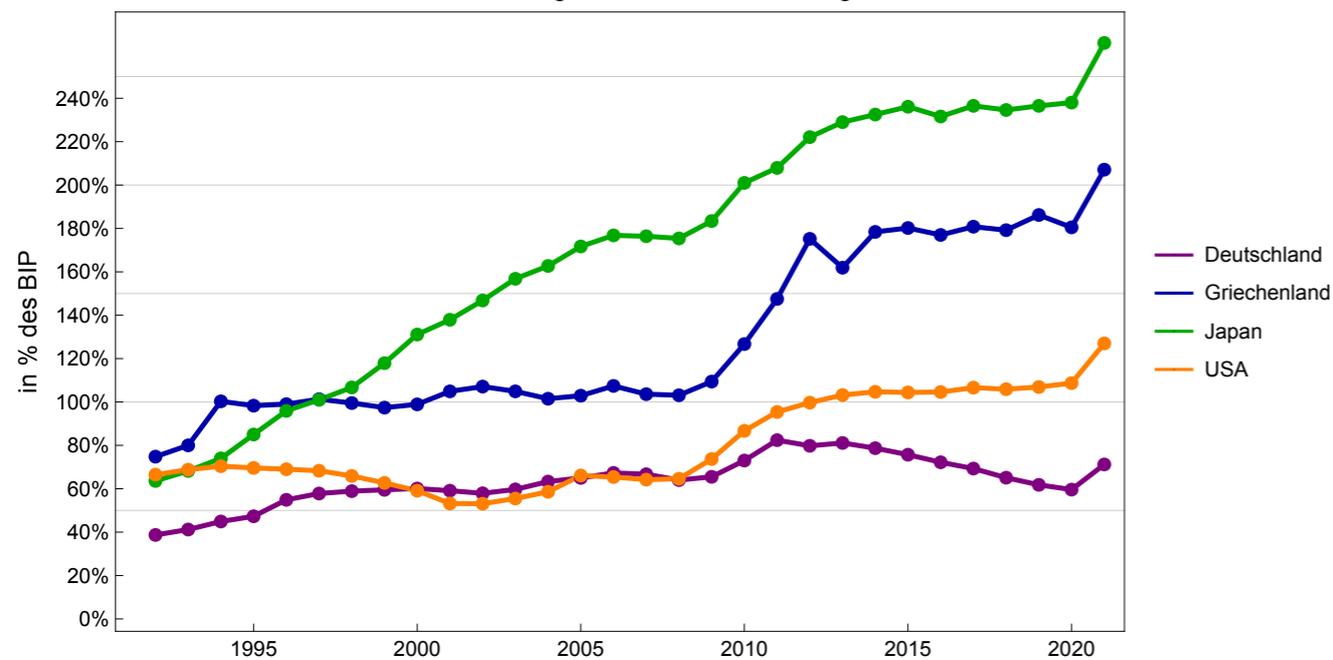
Staatsverschuldung im internationalen Vergleich



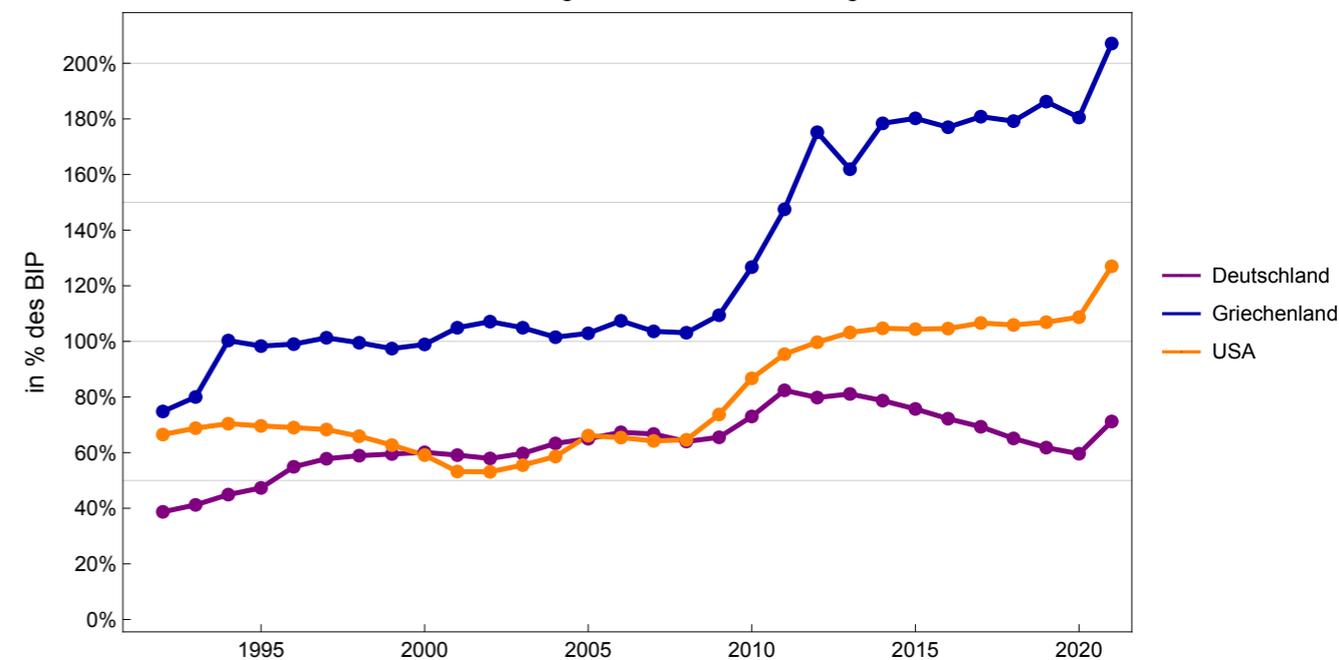
Gesucht: Spitzenreiter der Staatsverschuldung

- Internationale Vergleiche und die Wahl „passender“ Vergleichskandidaten
 - Diese prägen nicht uU nicht nur Minima und Maxima, sondern gesamte Skalierung.

Staatsverschuldung im internationalen Vergleich



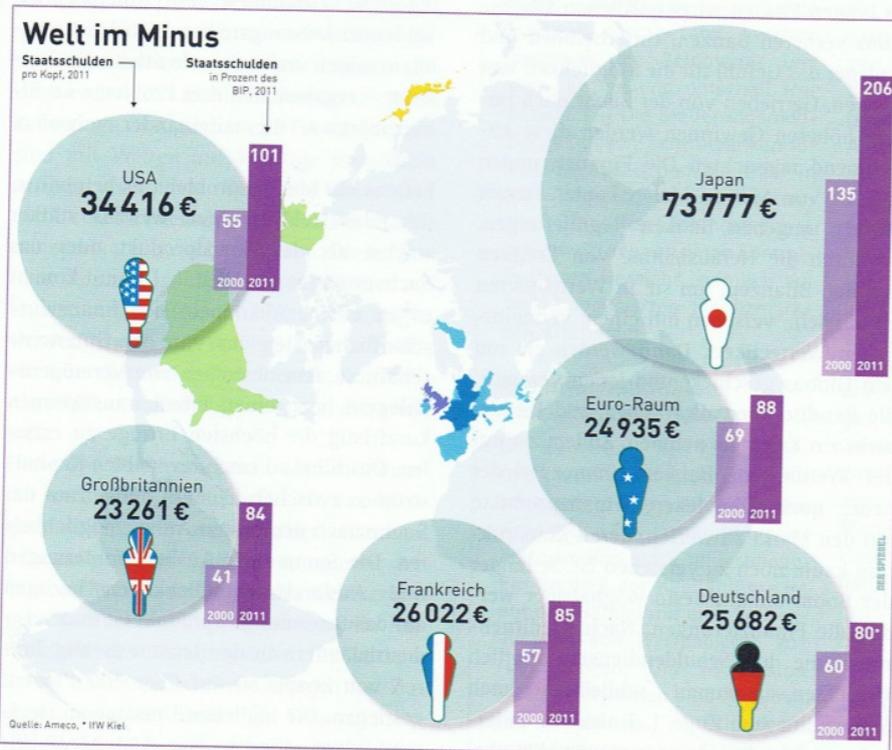
Staatsverschuldung im internationalen Vergleich



M 3 ● Die Aktienmärkte – von Krise zu Krise



M 4 ● Welt im Minus



Der Spiegel, 2.1.2012

Warum ist die untere Abbildung manipulativ?

Die Welt kann ohne interplanetarische Wirtschaftsbeziehungen niemals „im Minus“ sein

Aus der Saldenmechanik wissen wir:

Nettogeldvermögen des Privatsektors +
Nettogeldvermögen des Staates +
Nettogeldvermögen des Auslands = 0.

Die Summe der Geldvermögen aller Wirtschaftseinheiten weltweit ist immer Null.

Ergo: Den Staatsschulden stehen Privatvermögen gegenüber. Ohne Schuldaufbau kein Finanzvermögen.

Praktischer Test: Alle zuordenbaren Schulden und Guthaben sortieren.

Outcome: Lücke von 8% des globalen Finanzvermögens (ca. 5.800 Mil. €)

Grund: Verstecktes Privatvermögen in Steueroasen - „Schulden“ dokumentiert, „Guthaben“ nicht. (Zucman (2014): Steueroasen. Frankfurt/Suhrkamp.)

Wie hoch ist das mittlere Einkommen dieser fünf Personen?

Person 1	Person 2	Person 3	Person 4	Person 5
13000	14000	15000	16000	17000

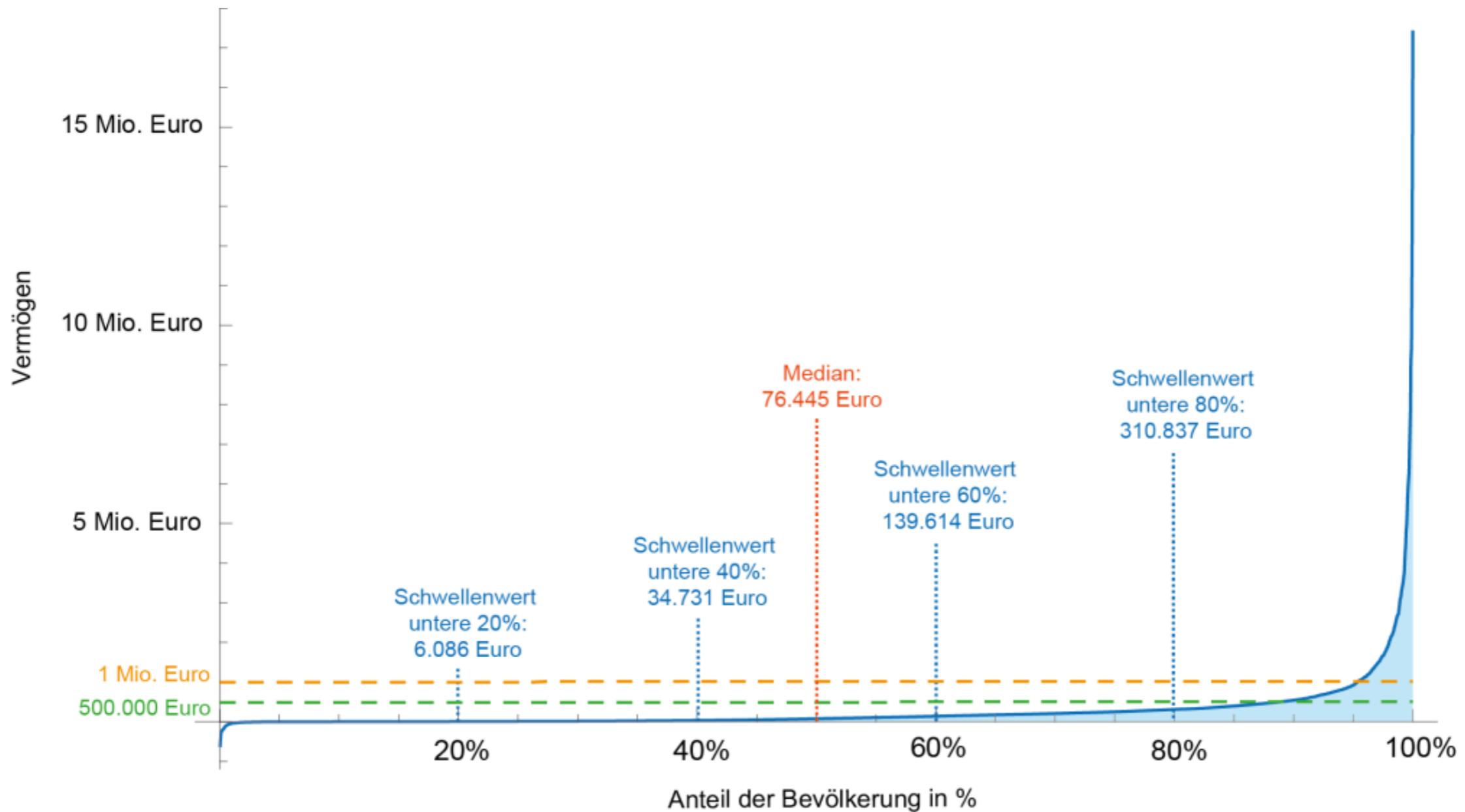
Wie hoch ist das mittlere Einkommen dieser fünf Personen?

Person 1	Person 2	Person 3	Person 4	Person 5
13000	14000	15000	16000	17000
13000	14000	15000	16000	442000

Wie hoch ist das „mittlere“ Einkommen dieser fünf Personen?

Person 1	Person 2	Person 3	Person 4	Person 5	Median	Arithm. Mittel
13000	14000	15000	16000	17000	15000	15000
13000	14000	15000	16000	442000	15000	90000

Vermögen und Einkommen: Mittelwert oder Median?



Verteilungskurve für Vermögen in Österreich; Eckerstorfer et al. (2013) http://media.arbeiterkammer.at/PDF/MWuG_Ausgabe_122.pdf

Resümee

“ Refusing to deal with numbers rarely serves the interests of the least well-off.”

Piketty (2014): Capital in the 21st century

Wiederholungsfragen zum Selbststudium

Wiederholungsfragen zum Selbststudium

- Welche allgemeinen Zielsetzungen und Anforderungen sind bei der Erstellung deskriptiver Statistiken zu beachten?
- Welche fünf klassischen Formen der deskriptiven Statistik wurden in diesem Abschnitt diskutiert? Welche drei davon sind besonders wichtig?
- Welche Gemeinsamkeit haben der „Spaghetti-Chart“ und der „Clutterplot“?
- Welche Darstellungsform wählen Sie um (a) die Entwicklung einer Zeitreihe, (b) die Verteilung von Körpergrößen in einer Schulklasse und (c) den Zusammenhang zwischen Lernmotivation und schulischer Leistung abzubilden?

Wiederholungsfragen zum Selbststudium

- Was ist mit dem „Stauchen von Achsen“ gemeint und was hat es mit graphischer Manipulation zu tun?
- Was ist mit dem suggestiven Kombinieren von Variablen gemeint? Welche beiden Beispiele dazu haben wir betrachtet?
- Wie sollten wir die Veränderung einer Wachstumsrate sprachlich beschreiben um nicht in potentielle Dramatisierung zu verfallen?

Wiederholungsfragen zum Selbststudium

- Welche Bezugsgrößen sind bei der Betrachtung der Entwicklung der nominellen Staatsausgaben besonders relevant?
- Was ist bei internationalen Vergleichen zu beachten? Inwiefern können diese manipulativ sein?
- Warum macht es konzeptionell wenig Sinn von einer „Welt im Minus“ zu sprechen?
- Warum kann es hilfreich sein Median und Mittelwert einer Kennzahl zu vergleichen?