

Computational Text Analysis and the Politics of Education

University of Marburg

Milan Thies

27.06.2025



www.eui.eu

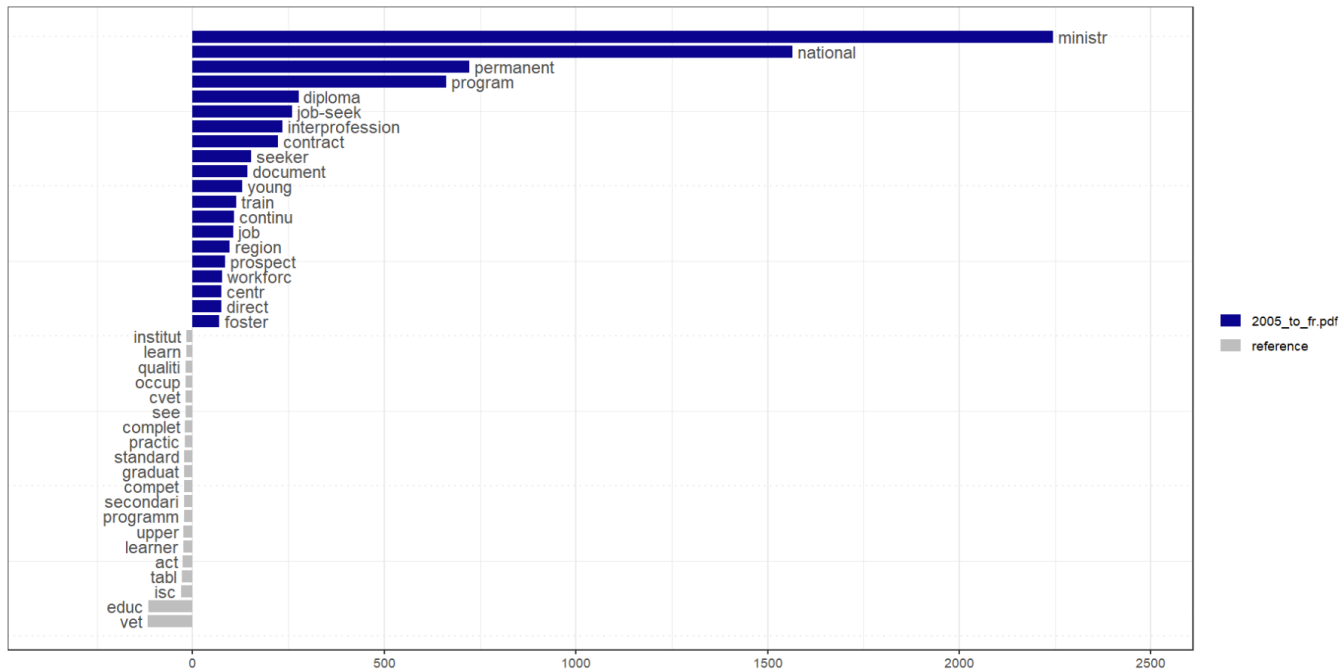


A Typical European Commission Press Conference

What is it about in one word or sentence?

<https://youtu.be/5CZRN0sbUVs>

What you will learn by tomorrow



Who am I?

- Researcher at the European University Institute in Florence
- I grew up in Marburg, happy to be back in town
- Worked in politics (e.g. writing speeches) and policy (writing government publications and inputs for legislation).
- Now I analyse the type of texts I produced before to better understand how education systems in Europe are changing.
- My expectations:
 - To help you to learn a new and useful method and maybe even raise some interest in education policy and politics
 - That you stick with me for these four days even if some days will seem tough. I promise it will be worth it in the end.




Examples from my own research: VET & Adult Ed - Early Childhood Education

Skill Formation, Natural Language Processing and the Analysis of Institutional Change: Towards a Common European Trajectory of Change?

Welfare states of industrialised economies exhibit a surprising level of resilience and stability (Esping-Andersen, 1996; Pierson, 2011). One especially stable institutional configuration of nationally unique welfare states are vocational education and training systems (Crouch et al., 1999; Hall & Soskice, 2001). Thelen (2004, p. 21) finds that “cross-national differences in vocational training regimes can be traced back to differences (...) in the early industrial period”. The consequences are *severe*. Vocational education and training, both as (post) secondary

Government and Opposition (2024), 1–22
doi:10.1017/gov.2024.25

 Government
& Opposition

ARTICLE

Conviction or Consent? Tracing the Influence of Coalition Partners on Family Policy under Centre-Right Ministers

Manuel Alvaríño  and Milan Thies 

Department of Social and Political Sciences, European University Institute, Fiesole, Italy

Corresponding author: Manuel Alvaríño; Email: manuel.alvarinovazquez@eui.eu

(Received 3 November 2023; revised 31 July 2024; accepted 9 August 2024)

Abstract

Many studies have analysed what could motivate centre-right governments to develop progressive family policies, given their historically traditionalist ideology. Updating classic institutionalist accounts, this article expands the focus beyond centre-right parties formally in charge. It argues that in coalition and minority governments, partisan veto players may act as agenda-setters, design policy reforms and successfully exert pressure to approve them through three mechanisms: agreements for government formation, conditions for government survival and bureaucratic continuity. Drawing on novel empirical data from interviews and document analysis, this article applies deductive process tracing

Who is this seminar for?

- Those that use text analysis but have not yet used computational methods or R to do it
- Those who have used R / quantda but want to revisit the basics
- Those who want to learn something about education politics and policy and how to research it
- Those who have no background in neither topic nor method but want to learn
- You are all welcome!

And now you: Who are you and what are your expectations for this seminar?

- Your name
- Your study programme
- Whether you have ever used R
- What you would like to learn during the seminar

Seminar Overview

Day 1 Theoretical foundations

- I. (Re)visiting the Policy Process & Texts as Traces of Political Conflict and Change
- II. The Politics of Education
- III. Qualitative Text Analysis

Day 2 Practical Foundations & First Analyses

- IV. A Newcomer's Guide to Computational Text Analysis
- V. Working with R: Introduction to R & Quanteda
- VI. Workflow & Descriptive Analyses

Seminar Overview

Day 3 Text-as-Data Methods

VII. Word Counts and Dictionaries

VIII. Supervised- and Unsupervised Machine Learning for Text Analysis

Day 4 Application

IX. In-Class Project: Applied CTA and the politics of education

X. Presentation, Reflection and Outlook

Practicalities

- 4 days. 09:30-13:00 and 14:00-17:30
- The seminar is for you, so do let me know if I can help you with anything
- You do some own data analysis on day 4. The first three days will prepare you for that small project. Big advantage: You can use the analysis you have done in the seminar after the seminar (e.g. master/bachelor thesis; Projektstudium)
- Are there any formalities that you need to fulfill for your study program?
- ChatGPT: Yes, you can use it for coding, but be aware of pitfalls.
 - University rules (What are you allowed/not allowed according to your university?)
 - Try to do it on your own first, only use ChatGPT to help out where you struggle
 - Wrong/bad advice is the rule rather than the exception. Specification is your friend (e.g. include “only use the quanteda package” in your prompt)
 - Try to understand what you are doing / never follow ChatGPT blindly

Questions before we start?

Session I.

(Re-)visiting the Policy Process & Texts as Traces of Political Conflict

(Re-)visiting the policy process: What is public policy?

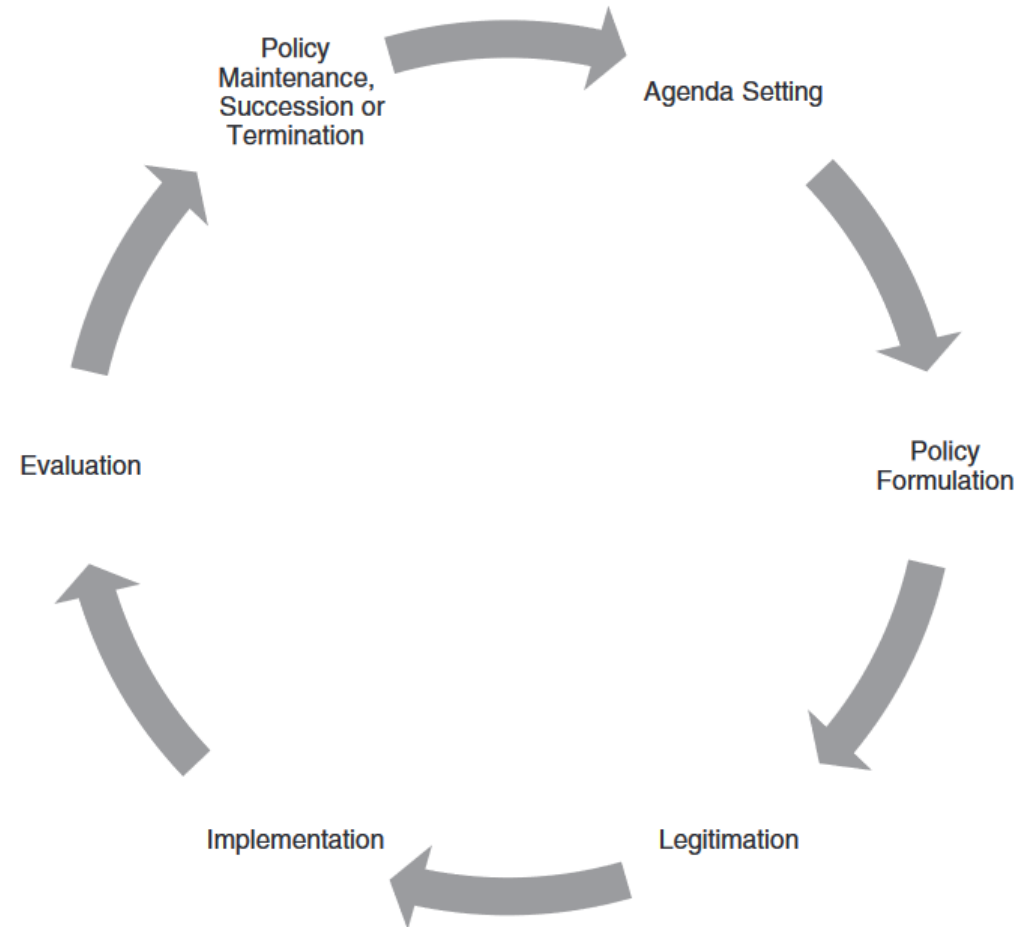
- No 'blind' methods seminar but applied to a policy field: The politics of education.
- We want to use the methods learned to understand how education systems are changing and why they are changing.
- Before we look at education politics let's clarify important terms: What is public policy and what is politics?

(Re-)visiting the policy process: What is public policy?

- “Public policy is what public officials within government, and by extension the citizens they represent, choose to do or not to do about public problems.” (Kraft & Furlong 2019, p.3)
- Politics is “the interaction between actors involved” (Hassel & Wegrich 2022, p.14) or “the political games being played between the various actors that are, or want to be, involved” (Hassel & Wegrich 2022, p.13) in policy-making.

(Re-)visiting the policy process

- **What then is policy-making?**
- **Policy process as simple but helpful model**
 1. Agenda setting: What is the problem?
 2. Policy formulation: What is the solution?
 3. Legitimation: Who supports the policy?
 4. Implementation: How can the policy be applied in the real world?
 5. Evaluation: How good does the policy work?
 6. Termination or maintenance: Should we keep, terminate or adopt the policy?



Policy-making beyond the policy process

- Understanding policy-making requires a more complex understanding of politics (= the interaction between actors).
- **‘The Dual Structure of Policy-Making’ (Hassel & Wegrich 2022)** divides the actors into an engine room and a superstructure. What do they mean by that?



Policy-making beyond the policy process

- Important lesson number one for applying text analysis but generally for any political science research:
Be aware that choosing to analyze certain data also entails a decision on which parts of policy-making will remain undisclosed.
- It matters, e.g. if you analyse parliamentary speeches or laws or think tank publications



Textual manifestation of political conflict

- How do political actors interact with each other in the policy process?
 - Express their positions and critique the positions of others via different media channels
 - Hold speeches defending positions or attacking the positions of others
 - Write manifestos and other textual documents to state their position
 - Mass movements chant their position on manifestations and write them on banners.
- Text is therefore not only at the center of politics but also the best way to learn about the positions of the different actors involved and the conflicts between them.

Example:

- From plenary debate of the law to strengthen lifelong learning and vocational education (04/2023)
- This is a speech by an opposition politician (CDU) responding to a proposal from an SPD minister.
- We can learn a lot about the positions of very different actors through this short text. Who can tell me a position of an actor from the text?

Gleichzeitig schaffen Sie ein neues Instrument. Was die Einführung des Qualifizierungsgeldes angeht, bleiben viele Fragezeichen. Denn es ist überhaupt nicht klar, welche Lücke in der Förderung damit geschlossen wird. Die Bewertungen des Deutschen Gewerkschaftsbunds und der IG Metall in ihren Stellungnahmen in der Anhörung, Herr Minister, waren geradezu vernichtend.

(Hermann Gröhe [CDU/CSU]: Aha! So ist es!) (D)

Die sagen: Das ist nicht nur zu komplex, sondern das ist für die betriebliche Praxis einfach auch unattraktiver als das, was wir schon haben. Das wird niemand in Anspruch nehmen. – Lieber Herr Minister, wenn Ihnen sogar die Gewerkschaften so ins Gewissen reden, dann müssen Sie nacharbeiten.

(Beifall bei der CDU/CSU sowie bei Abgeordneten der LINKEN)

Spannend ist auch, was nicht im Gesetzentwurf steht, so wie er im Parlament angekommen ist; denn hier ist sich die Ampel mal wieder nicht einig. Die vorgesehene Bildungszeit ist nach Intervention des Finanzministers wieder rausgeflogen wegen unberechenbarer Finanzauswirkungen.

Textual manifestations of political change

- Texts do not only capture politics (actor interactions) but also policy (result of interactions).
- Important types of documents include e.g. laws, government publications, policy descriptions
- Questions to ask: What is changing? What problem does this address?
- Example: Screenshot of European Skills Agenda Communication.

What are the problems this aims to address? What is done?

Action 5: Rolling out the European Universities initiative and upskilling scientists

To roll out the European Universities, the Commission, in close cooperation with the stakeholders and the Member States, will:

- engage in the full rollout of the European Universities initiative under the Erasmus programme (2021–2027) and Horizon Europe, including by removing obstacles to effective and deeper transnational cooperation between higher education institutions and deepening the cooperation with economic operators, in particular to foster the twin transitions. European Universities will set standards for the transformation of higher education institutions across the European Education Area and the European Research Area, also making lifelong learning and talent circulation a reality.*
- explore options stemming from their research and innovation dimension to help remove obstacles to effective transnational cooperation between higher education institutions, drawing on the lessons learnt during the pilot calls under Erasmus+ and Horizon 2020. The Commission will identify areas of support for Member State action, explore a concrete approach for a “European degree” and the feasibility of a European University statute (to tackle cross-border legal issues) and for a European Recognition and Quality Assurance System.*

Session II.

The Politics of Education

Segments of Education Systems

- Early childhood education
- Primary and secondary education (= school education)
- Vocational education
- Higher education
- Lifelong learning (= adult education)

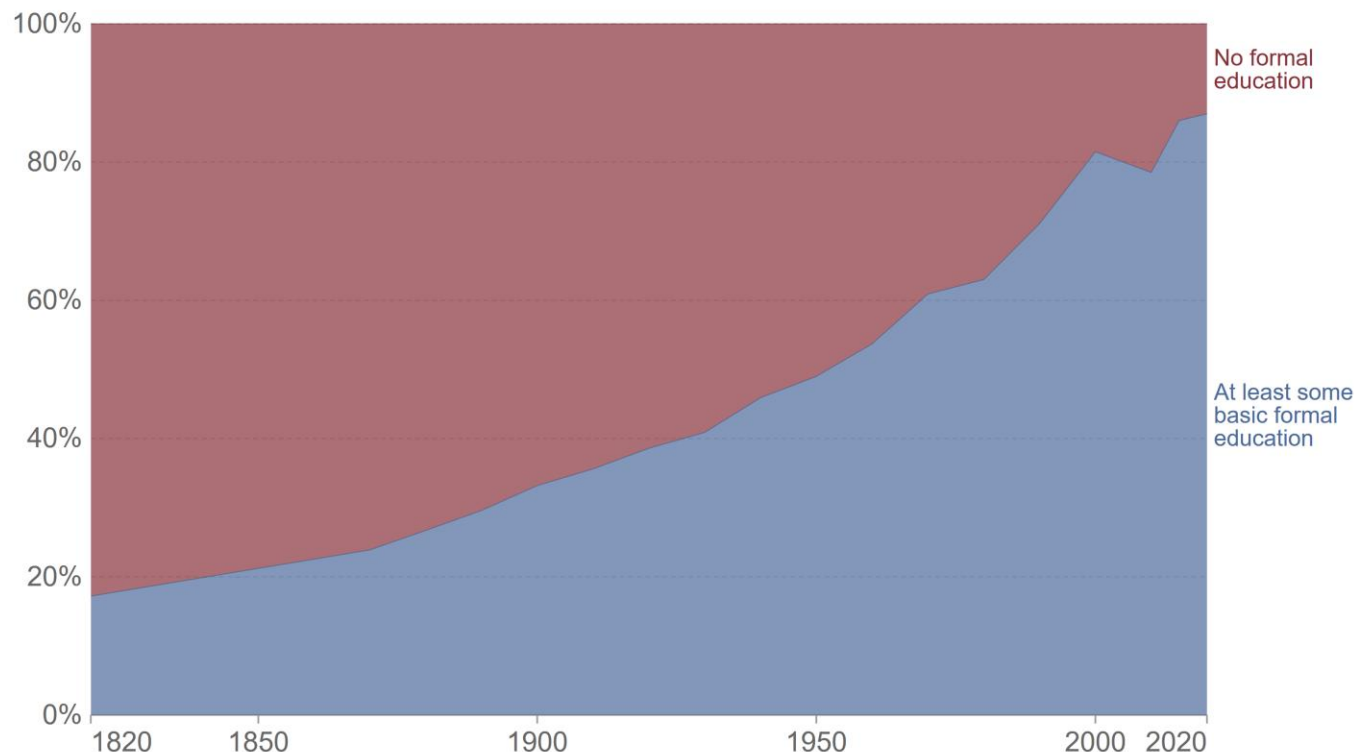
Education systems over the long run

- **Education revolution** transformed society in just over 1,5 centuries
 1. Formal education shifted from a privilege of view to a “mandatory essential for everyone” (Baker 2014)
 2. Global spread of certain ideas and norms (culture). John Meyer: Consequence is a world society where culture, economy and political systems converge

Share of the world's population with formal basic education

Share of people aged 15 or older who received some kind of formal primary¹, secondary², or tertiary³ education.

Our World
in Data



Data source: Our World in Data based on Wittgenstein Centre and van Zanden, J. et al. (2014) via OECD

Note: Formal education is primary¹ or higher. Data for the years before 2015 are estimates, while data from 2015 onwards are projections

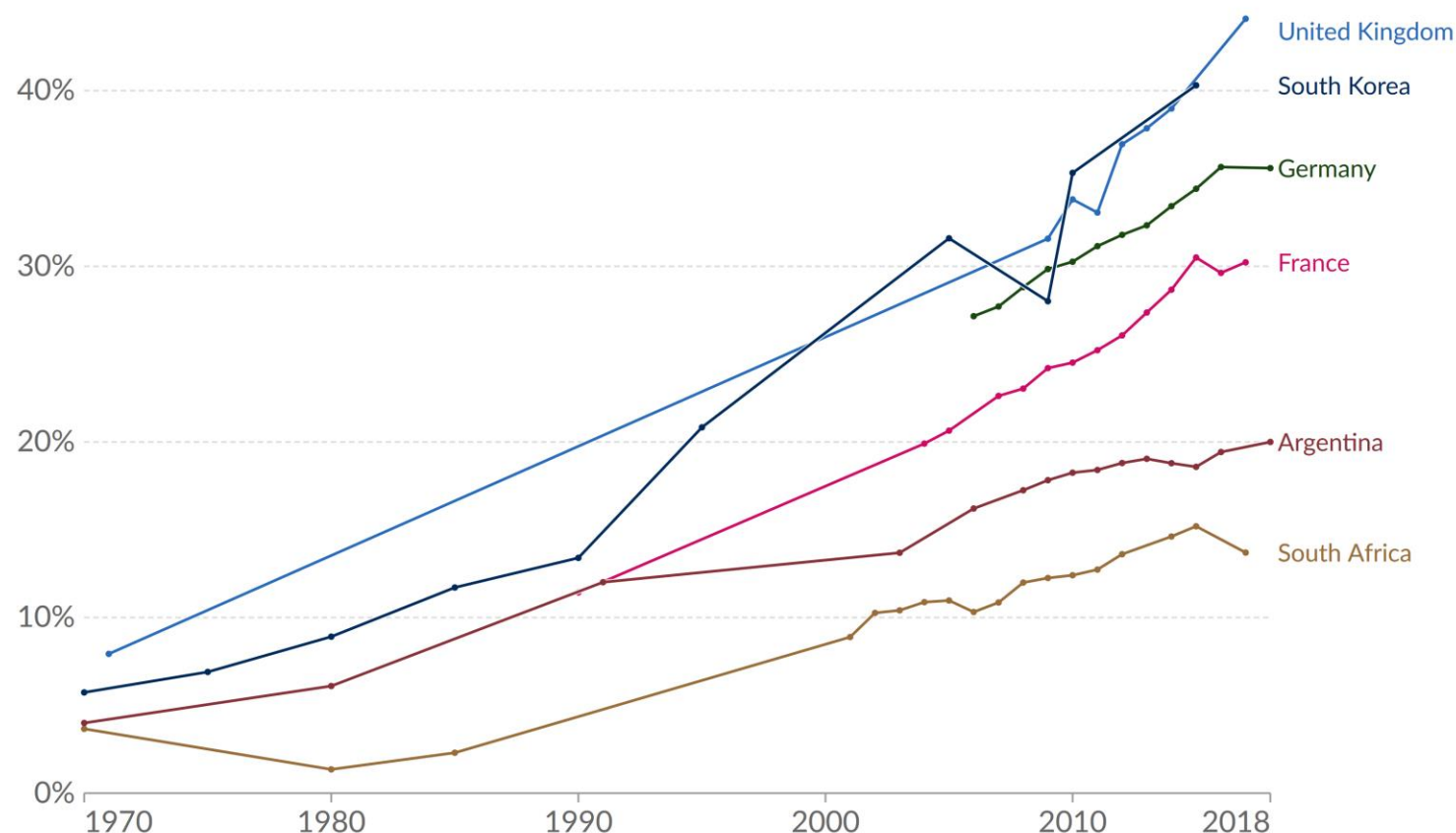
OurWorldInData.org/global-education | CC BY

- In reality: No full convergence but different trends of rising participation
- Key trend in industrialized societies: Transition towards knowledge economies.
- Fundament thereof: Massive expansion of higher education

Share of the population with post-secondary education, 1970 to 2018

Percentage of the population aged 25 and over who have completed post-secondary education (ISCED level 4 or higher).

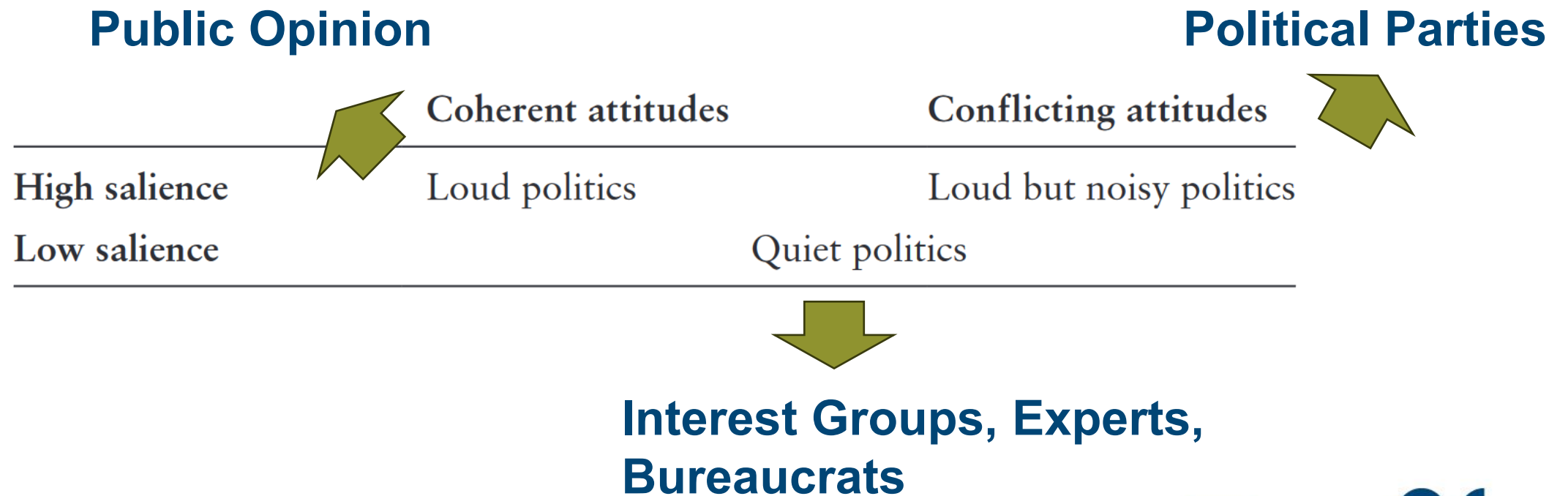
Our World
in Data



Data source: UNESCO via World Bank

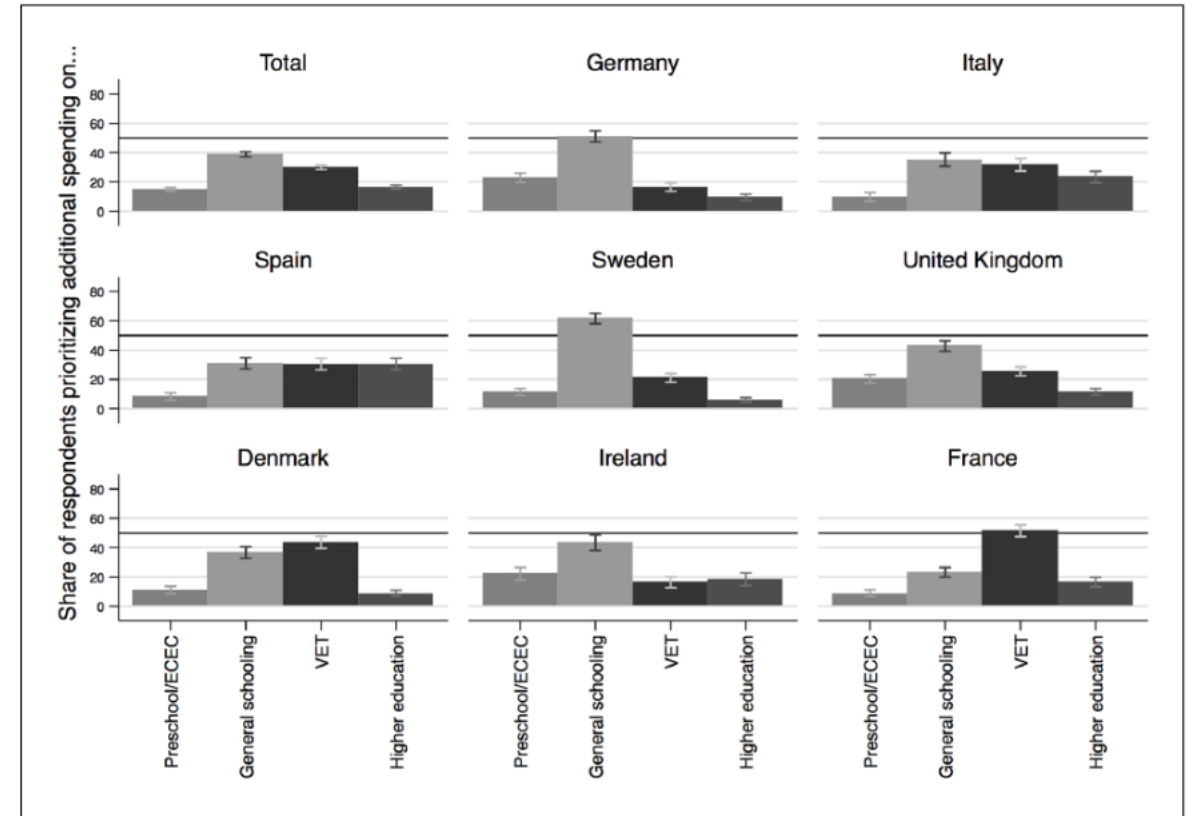
OurWorldInData.org/global-education | CC BY

What determines education politics? (Busemeyer et al. 2022)



Education Politics & Public Opinion

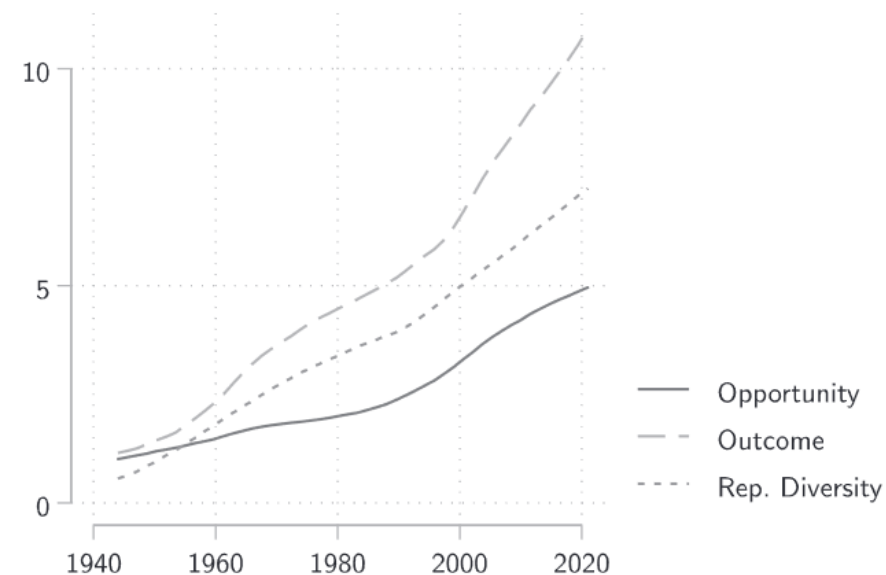
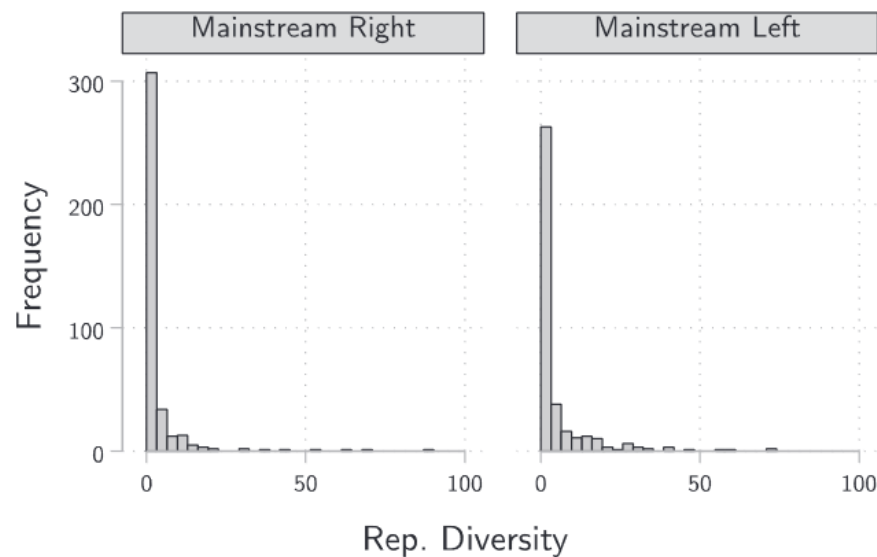
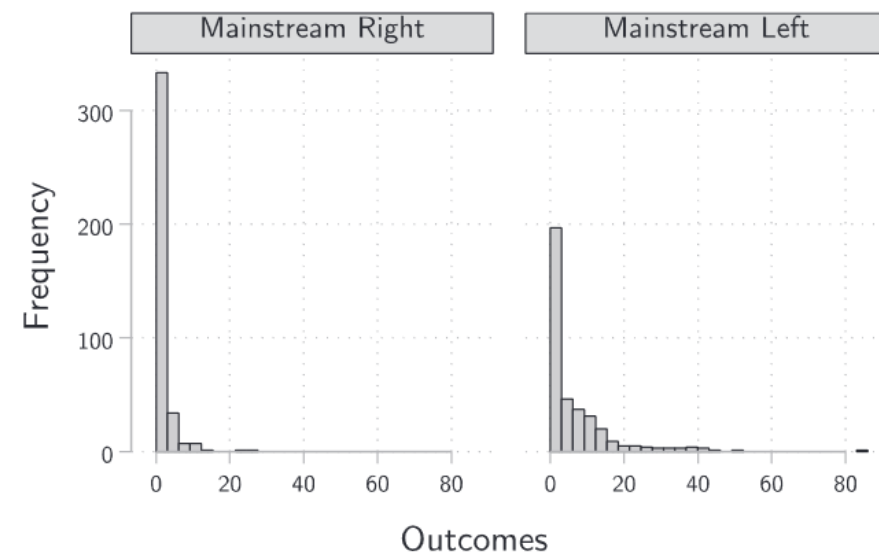
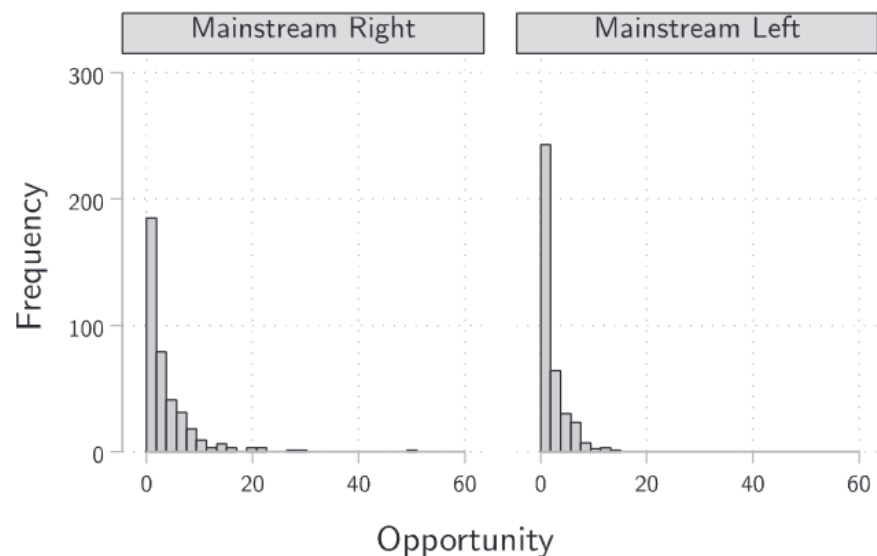
- Loud politics: Public largely agrees on position and has a high priority for it.
- In Germany: Majority in favor of additional spending on general schooling, strong priority over other areas
- However, we do not see a prioritization of spending matching public opinion
- Therefore, educational expansion not clearly loud politics



The Partisan Politics of Education

- Should matter most when politics is “loud but noisy” (= high salience, divided public opinion)
- Educational expansion is generally a consensual issue supported both by the left and right (Jakobi, 2011)
- Even the goal of using education to increase equality has become consensual. However, what concept of equality is meant differs between parties (Gingrich & Giudici 2022)
 - Equality of opportunity (procedural fairness, meritocracy)
 - Equality of outcome (decreasing social inequality outcomes)
 - Equality of representation (decreasing between-group differences)

- Gingrich & Giudici (2022): Frequency of term mentions associated with different concepts of equality in education
- Source: Party manifestos
- What do we see here?



Interest groups and education policy

- Depends on educational sector. In adult education and VET private actors are often providers themselves.
- Example: German VET reform of 2005 introduced a second track of school-based VET alongside the traditional dual apprenticeships which did not provide sufficient apprenticeships.
- Employers & labour unions tried to avoid “intrusion of state”. While they ultimately failed, they managed to keep the state from introducing a larger reform and school-based VET in Germany remains the exception rather than the norm.
- **Take-away: Interest group politics matters, but often in interaction with other drivers of public policy.**

So, who matters in education politics?

- Reality is more complicated than models such as the one developed by Busemeyer et al. (2022) which still is a good point to start (and good hypotheses for any academic papers, master theses,... you might write)
- This theory of loud, loud but noisy and quite politics is very useful for text-as-data analyses: We can compare what actors demand and what changes are adopted through computational text analysis.

Conflicts in Education Politics: Privatization vs State Provision

- While education is generally considered a public task, much is provided by private actors (especially VET and adult education)
- What do you think is the consequence of increasing public or private provision of education?

Privatization vs State Provision

- The general assumption: The more involvement of private actors, the more economic efficiency (labor market match) is achieved. More involvement of the state in turn can benefit social inclusiveness.
- Past decades have seen a trend of liberalization leading to more privatization, deregulation and reliance on markets. However, more recent research finds the opposite trends in recent years.

Allmendinger (1989):
Stratification and standardization
in schools (top) and vocational
education (bottom)

Stratification vs inclusion

How much standardization?

- Standardization: The rules and procedures that assure employers of what graduates know
- Stratification: The proportion of the population that will achieve the highest number of (school) years provided by a system

		Stratification	
		high	low
Standardization	high	FRG NORWAY (secondary)	
	low		USA NORWAY (primary)

		Stratification	
		high	low
Standardization	high		NORWAY FRG
	low	USA (on-the-job)	

Stratification vs inclusion

Allmendinger (1989):
Stratification and standardization
in schools (top) and vocational
education (bottom)

- Highly politicized question!
- German school system is highly stratified. Ever since the 1970s: Debate where left-wing parties have promoted *comprehensive, inclusive schools* and conservative parties were in favor of the *traditional segmented school system*.

		Stratification	
		high	low
Standardization	high	FRG NORWAY (secondary)	
	low		USA NORWAY (primary)

		Stratification	
		high	low
Standardization	high		NORWAY FRG
	low	USA (on-the-job)	

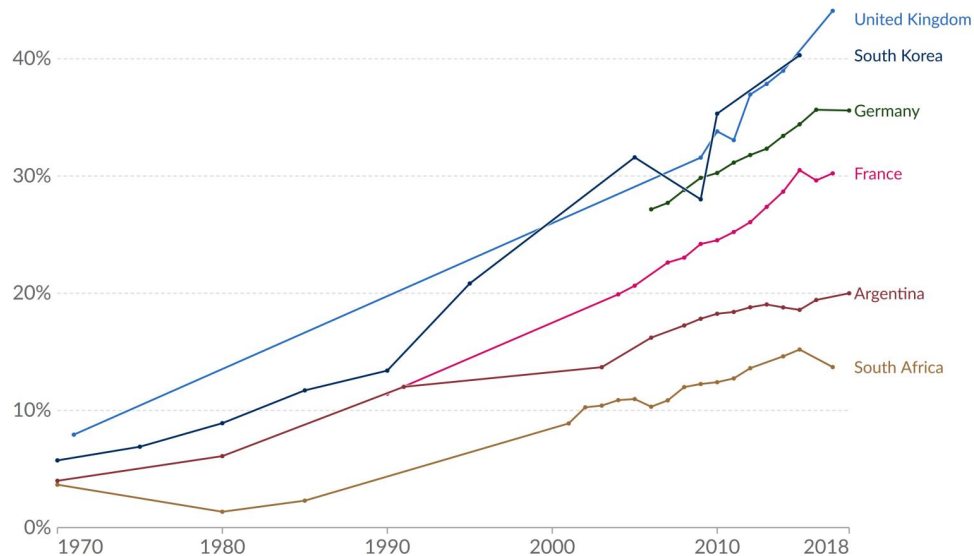
Early Childhood-, General-, Higher- or Vocational Education?

There are different priorities across countries

Share of the population with post-secondary education, 1970 to 2018

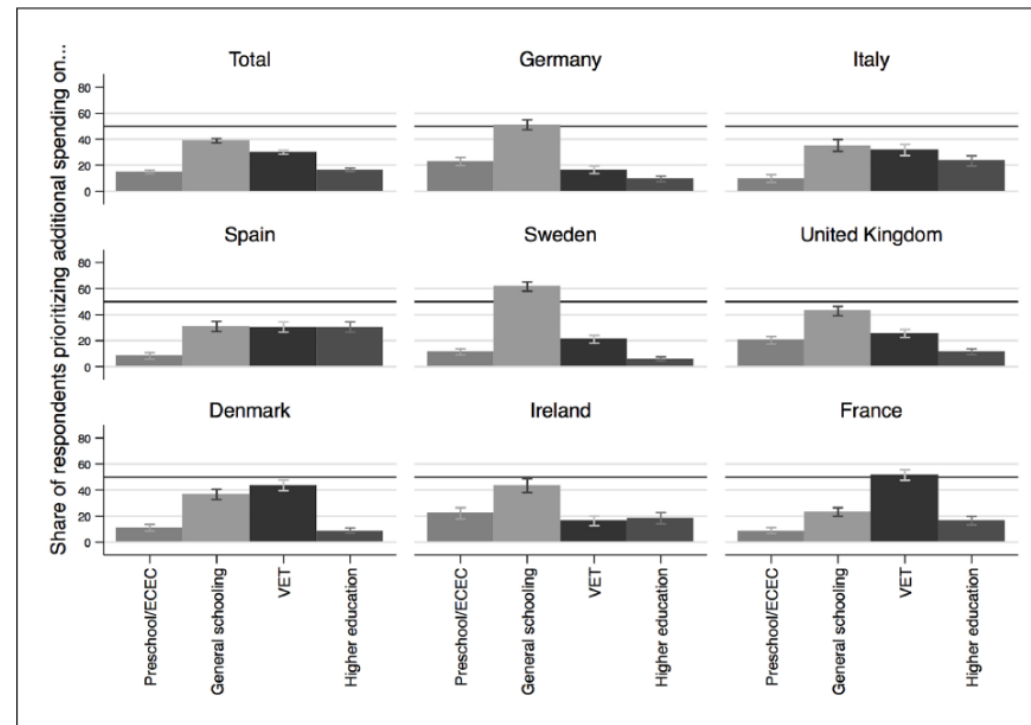
Percentage of the population aged 25 and over who have completed post-secondary education (ISCED level 4 or higher).

Our World
in Data



Data source: UNESCO via World Bank

OurWorldInData.org/global-education | CC BY



Education as enabler of transformation

- Newest development is the role of education in facilitating (the digital and green) transformation
- Example: European Skills Agenda. Here first paragraph of it.

We live in a time of transitions. The twin green and digital transitions are reshaping the way we live, work and interact. The EU's move to a resource-efficient, circular, digitised and climate neutral economy and the wide deployment of artificial intelligence and robotics are expected to create new jobs¹ while other jobs will change or even disappear. Demographic change will require Europe to draw on all of its talents and diversity. At the same time, it will also generate new job opportunities in the silver and care economies. These transitions show the need for **an unparalleled shift in skill sets** to reap their full potential.

And now you:

1. Think about one education policy issue you care about
2. Tell you neighbor about it and discuss:
 - What is the problem and what might be an approach to tackle it?
 - Who are the actors who matter? What might be their positions??
 - Is it quiet, loud, or loud-but noisy politics?
 - Does it connect to any of the conflicts I outlined?
3. Present the issue your neighbor cares about

<https://www.youtube.com/watch?v=ojNkeyxQmGA>

Session III.

Qualitative Text Analysis

Qualitative Text Analysis: Content Analysis

- One of most used methods in qualitative research (Kuckartz 2019)
- Transparent, reliable and easy to learn
- Basic idea: Making sense of content of texts by assigning codes and categories to text
- At core it is about category building, either inductively or deductively
- Important: Qualitative content analysis can also be computational, does however not belong to the field of 'text as data' which is purely quantitative
- Can be complementary to quantitative analysis to validate more automated methods of text analysis

ARTICLE

Conviction or Consent? Tracing the Influence of Coalition Partners on Family Policy under Centre-Right Ministers

Manuel Alvarino and Milan Thies

Department of Social and Political Sciences, European University Institute, Fiesole, Italy

Corresponding author: Manuel Alvarino; Email: manuel.alvarinovazquez@eui.eu

(Received 3 November 2023; revised 31 July 2024; accepted 9 August 2024)

Abstract

Many studies have analysed what could motivate centre-right governments to develop progressive family policies, given their historically traditionalist ideology. Updating classic institutionalist accounts, this article expands the focus beyond centre-right parties formally in charge. It argues that in coalition and minority governments, partisan veto players may act as agenda-setters, design policy reforms and successfully exert pressure

50

P2: Ja. Die beiden großen familienpolitischen Vorhaben aus der damaligen Zeit, die Einführung des Elterngeldes genauso wie die Einführung des Kitarechtsanspruchs, die gingen im Wesentlichen auf die Initiative von SPD Politikerinnen zurück. Das Elterngeld hat Renate Schmidt in Deutschland populär gemacht. Das hatte ja skandinavische Vorbilder. Aber sie war diejenige, die es nach Deutschland transportiert hat und dann so umgewandelt hat, dass es auf deutsche Verhältnisse passt. **Das war ihre Idee, das war ihr Konzept, und das war ihr politisches Vorhaben.** Und beim Kita Rechtsanspruch war es eigentlich genauso, also der Kita Rechtsanspruch, so wie er auch nachher umgesetzt worden ist mit den Eckpunkten. Der ist im Wesentlichen in einer Arbeitsgruppe entstanden aus SPD Fraktion und SPD Parteivorstand. Da waren alle möglichen Leute beteiligt, das waren Bundestags, SPD, Bundestagsabgeordnete, Landespolitikerinnen, Kommunalpolitiker, Partei, stellvertretende Parteivorsitzende. Also das war ein SPD Konzept. **Die Frau von der Leyen hat das Elterngeld so umgesetzt wie es im Koalitionsvertrag stand.** Ja, sie war diejenige, die in der Verantwortung war. Sie war da einfach klug. Und da hat sie dann die Karnevalsjaahr 2700 nachrichtenarme Zeit und wird das Thema gesetzt, ganz, ganz prominent für jedes dritte Kind unter drei brauchen jetzt einen Kitaplatz. Das war ungefähr zu der Zeit, wo die SPD mit ihrem

Agenda Setting + 2005 Election
Coalition Formation

Coalition Formation

Coalition Formation
Coalition Management + Parliamentar

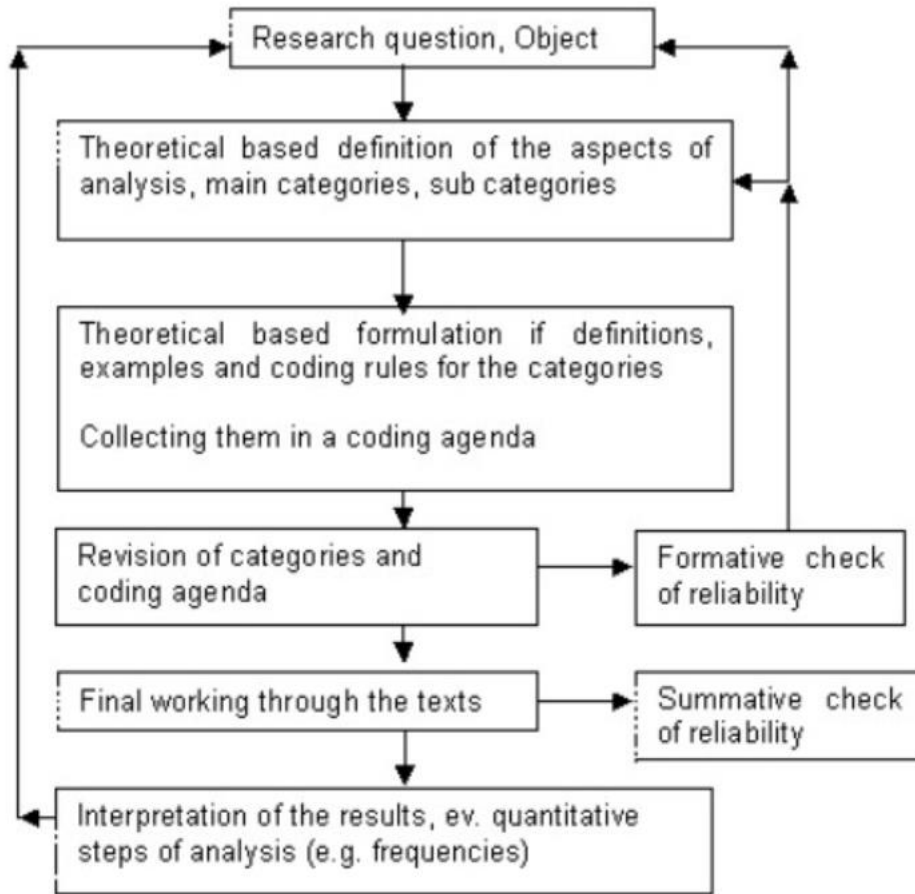
- Paper on early childhood education reform in GER
- Question: Did the centre-right party CDU introduce a progressive reform out of conviction or did they have to consent against their own ideology? Why?
- Codes represent mechanisms of influence:
 - Agenda setting
 - Electoral campaign
 - Coalition formation
 - Coalition management and parliamentary process

MaxQDA as example of software to help with content analysis

The screenshot displays the MaxQDA software interface, which is used for qualitative data analysis. The interface is divided into several panels:

- Document Browser (Top Left):** Shows a list of documents. The document '210412' is selected, with a count of 57 paragraphs. Other documents listed include '210317', '21051', '21033', '21040', '21042', '210416', and '21042'.
- Sets (Middle Left):** Shows a hierarchical structure of sets. The 'Germany' set is expanded, showing 'France' with a count of 185. Other sets include 'More features of Administrative ...' (6), 'Europeanization' (4), 'Economic Development' (7), 'Technological Change and Digit...' (11), 'Political Calculation' (16), 'Prior Laws and Agreements' (46), 'Dist. of Responsibility: National ...' (38), 'Dist. of Responsibility: National ...' (35), 'Dist. of Responsibility: State, So...' (49), 'Reform Content' (0), 'Reform of Control: Other than C...' (48), and 'Reform of Provision' (21).
- Document Viewer (Center):** Displays the content of the selected document '210412'. The text is in German and discusses vocational training and the role of the state. A yellow highlight is visible on the text: 'mit einem Blick darauf: was hat sich dann durch die nationale Weiterbildungsstrategie verändert?'. Below the text, there is a vertical timeline with markers for '11', '12', and '13'.
- Code Browser (Bottom Left):** Shows a list of codes. The code 'Reform of Provision' is selected, with a count of 21. Other codes include 'Reform of Control: Other than C...', 'Reform Content', 'Dist. of Responsibility: State, So...', 'Dist. of Responsibility: National ...', 'Dist. of Responsibility: National ...', 'Prior Laws and Agreements', 'Political Calculation', 'Technological Change and Digit...', 'Economic Development', 'Europeanization', and 'More features of Administrative ...'.
- Memo 5 (Right Panel):** A yellow box containing the text: 'Perspektive of BDA on what the state and what the firms should do: Social partners together monitor skill demands and then adopt continuing vocational training. State sets framework and finances social needs. According to BDA it should leave the rest up to firms and the continuing education market.'

Deductive Workflow



- Step1:
 - What education sectors are politicians talking about in parliament?
 - Source: Parliamentary speeches
- Step 2 and 3: VET, school-education, higher education, adult education, early childhood education
- Step 4: Work through speeches
- Step 5: Interpret results

And now you!

- Download “Content Analysis Coalition Treaties” from ILIAS or my website (<https://milanthies.github.io/teaching>)
- It contains translated paragraphs from German coalition treaties.
- You have to code what goal the treaty content fulfils (if any).
- Do not discuss with your neighbour at this point. There is no wrong or right, but it is interesting to see how many of you agree.

Introduction to Skill Formation Systems & Research Critique

- Let's apply your knowledge of structured content analysis
- We will read Graf & Marques (2022) and then discuss how they apply content analysis
- The reading connects to several debates of today: On different education sectors, on inclusiveness and efficiency of education systems and on convergence

Introduction to Skill Formation Systems & Research Critique

- Skill formation systems are mostly about vocational education and training
- 4 dominant types
- Collective systems: Balancing inclusiveness and efficiency?
- Skill formation systems are said to be very stable ever since the industrial revolution

Table 1 – Typology of Skill Formation Systems; own illustration based on Busemeyer & Trampusch (2011)

Skill formation system	Public Commitment	Involvement of Firms
Liberal	Low	Low
Segmentalist	Low	High
Statist	High	Low
Collective	High	High

Introduction to Skill Formation Systems & Research Critique

- What is the research question Graf & Marques (2022) want to answer?
- What text data do they use?
- Who created the text data?
- How do they analyse the text?
- What do they find?
- Do you think this is a good research design? Is their analysis convincing?

Did everyone manage to install R, R-Studio and the package quanteda?