RESEARCH ARTICLE

Explaining differences in research utilization in evidence-based government ministries

Jesper Dahl Kelstrup[1] · Jonas Videbæk Jørgensen[1]

© The Author(s) 2024

Abstract

Studies of evidence-based policy have found that research often fails to influence policy-making and identify a number of barriers to research utilization. Less is known about what public administrations do to overcome such barriers. The article draws on a content analysis of 1,159 documents and 13 qualitative interviews to compare how and why evidence standards affect research utilization in two Danish ministries with available evidence, policy analytical capacity, and broad political agreement on key policy goals. The article finds support for the proposition that more exclusive evidence standards in ministries will lead to higher levels of research utilization by showing that average levels of research utilization are higher in the Ministry of Employment than in the Ministry of Children and Education in the period 2016–2021. In active employment policy the adoption an evidence hierarchy and the accumulating evidence in a knowledge bank has interacted with stakeholder support and a continued coordination with the Ministry of Finance to provide economic incentives for policy-makers to adopt evidence-based policies thus stimulating research utilization. Evidence for public education policy, by contrast, has been more contested and the Ministry of Children of Education retains inclusive evidence standards in an attempt to integrate evidencebased and practical knowledge from stakeholders, which has led to lower average levels of utilization in the period.

✉ Jesper Dahl Kelstrup
kelstrup@ruc.dk

Jonas Videbæk Jørgensen
videbaek@ruc.dk

[1]   Department of Social Sciences and Business, Roskilde University, Roskilde, Denmark

Introduction

The idea of basing policies on the best available evidence has spread from medicine to several policy fields in the twentieth century (Baron, 2018; Head, 2015). Today, evidence is an important resource for governments in most modern democracies to identify effective solutions to policy problems (Adam et al., 2018, 2019; Davies et al., 1999; Parkhurst, 2017; Parsons, 2004; Sanderson, 2002). Policy scholars have examined the mobilization of evidence by research agencies (Powell et al., 2018), the capacities of governments to deliver policy analysis and advice (Howlett, 2009; Migone & Howlett, 2022), and the use of evidence in various decision-making bodies, including government agencies (Landry et al., 2003), state agencies (Jennings & Hall, 2012), and legislatures (Geddes, 2021). Yet, many studies find that research often fails to influence policy decisions and that research utilization varies across policy areas and contexts (Boswell, 2009; Landry et al., 2003; Lester, 1993; Lindblom & Cohen, 1979; Oliver et al., 2014; Weiss et al., 2008). Low levels of research utilization have been attributed to the shifting preferences of policy-makers (Majone, 1989), value conflicts over policy aims (Boswell, 2009), and technical barriers to the use of evidence, including poor timing between research and policy, lack of robust impact evaluations, and a poor fit between evidence-based findings and policy priorities (Cairney, 2016; Capano & Malandrino, 2022; Nutley et al., 2007; Weiss, 1995). The continuous importance of political and practical concerns has led some scholars to abandon the notion of "evidence-based policy" in favor of the more modest phrase "evidence-informed policy" (Bundi & Pattyn, 2022; Head, 2015; Nutley et al., 2007). Nevertheless, the extent to which evidence is used in public administrations remains an empirical question that is affected both by the level of policy analytical capacity and conflict in different policy subsystems (Howlett, 2009; Jennings & Hall, 2012).

The article adds to theory and research by comparing how and why evidence standards affect research utilization in two ministries with available evidence, similar policy analytical capacities and broad political agreement on key policy aims. The article relies on a most-similar case study design and a combination of content analysis of documents ($N=1{,}159$) and interviews with civil servants and relevant external stakeholders ($N=13$) to compare and explain research utilization in the Ministry of Employment and the Ministry of Children and Education from 2016 to 2021. By addressing the effect of evidence standards on research utilization in policy decisions, the article advances the study of research utilization and evidence-based policy as a limited number of contributions have measured levels of research utilization over time or across different public administrations (cf. however, Jennings & Hall, 2012; Landry et al., 2003; Newman et al., 2016).

The analyses capture both smaller and larger policy decisions, such as the 2014 Public School Reform, which has had a significant impact on the Danish education system, and a series of labor market reforms adopted in the period under investigation. While both ministries have invested in building policy analytical capacity to accumulate and consider evidence in the policy process, the content analysis finds higher average levels of research utilization and higher utilization of studies from the top of the evidence hierarchy in the Ministry of Employment than in the Ministry of Children and Education. Interview respondents attribute this variation in research utilization to internal efforts in the Ministry of Employment to adopt specific evidence standards and its investment in creating a knowledge bank for rating policy effects based on evidence. In addition, the Ministry of Employ-

ment has had a continuous dialogue and coordination with the Ministry of Finance to agree on principles for utilizing evidence in budget forecasting to estimate future economic gains with general support from stakeholders in the policy subsystem. By contrast, the Ministry of Children and Education has had a more conflicting relationship with stakeholders about letting evidence shape policy reforms. Even though the central ministry department has recently invested resources in a knowledge bank, there is not yet sufficient evidentiary basis to persuade the Ministry of Finance to include derived economic effects of public school policies in economic modeling, thus contributing to lower utilization levels.

Evidence-based policy and the politicization of evidence

The so-called "evidence movement" and its emphasis on rational decision-making has been an important driver of evidence-based policy, promoting that research is used to find policy solutions that have been documented to "work" across policy domains (Baron, 2018; Bundi & Pattyn, 2022; Head, 2015; Howlett, 2009; Parsons, 2004; Sanderson, 2002). Such "what works" evidence is typically produced following an evidence hierarchy, which values knowledge based on the ability of research designs to systematically eliminate bias when determining the causal effects of policies (Evans, 2003; Nutley et al., 2013; Petticrew & Roberts, 2003). The idea of using evidence to foster utilization based on policy effects as well as costs is also central to the rational model of decision-making because it can underpin policy goals of enhancing the cost-effectiveness of policies (Davies et al., 1999; Greany & Brown, 2017; Oliver, 2022). An additional appeal of evidence-based policy decisions taken at central levels of government, is that even small policy changes can have substantial aggregate effects if applied at scale, e.g., to whole populations (Baron, 2018). Evidence can therefore help identify effects that are not immediately visible to policy-makers but have effects on desired policy aims. On this basis, evidence-based policy has moved beyond the idea that politics and science are separate communities with different knowledge ideals and incentives and has spread to more policy domains over time including employment and education (Baron, 2018; Caplan et al., 1975; Caplan, 1979; Head, 2015).

Existing evidence-based policy studies, however, have identified barriers to research utilization and voiced criticisms of using a rational model of decision-making in public policy (Capano & Malandrino, 2022; Fischer, 2021; Lester, 1993; Newman et al., 2017). The politicization of evidence and the challenge that policymakers may cherry-pick evidence to suit their preferences have been emphasized as adverse uses of evidence (Weiss, 1979; Oliver, 2022; Parkhurst, 2016, 2017). In this view, evidence works as a reservoir that can be drawn upon for symbolic or strategic reasons when politicians desire to justify decisions (Daviter, 2015; Majone, 1989; Mosley & Gibson, 2017) or to pre-empt public criticism of government policies (Boswell, 2009). Several studies have thus found that research production and utilization are shaped by actor preferences, especially those of political decision-makers (Boswell, 2009; Fobé & Brans, 2013; Sanderson, 2002). "Policy-based evidence" has been suggested as a more appropriate term for evidence that political actors use to underpin pre-determined policy positions (Sanderson, 2011). If policy aims are changing or are disputed, the evidence might become a "moving target", as previously collected evidence can become irrelevant, e.g. with a change of government. Moreover, continued external politicization of evidence can interact with policy contestation and become a barrier to research utilization, as ministries may refrain from focusing their capacity on evidence collection and utilization

in light of policy conflict (Jennings & Hall, 2012). Even in cases where there are low levels of conflict, there is a gap in understanding mechanisms that link capacity to research utilization. This article focuses on the effect of policy analytical capacity and evidence standards on research utilization to address this gap.

Explaining the effect of evidence standards on research utilization

Notwithstanding politicization and policy conflict as barriers to evidence-based policy, there are drivers, pressures, and opportunities for public administrations to increase their policy analytical capacity and utilize evidence systematically and over time (Carpenter & Krause, 2015; Christensen, 2022; Howlett, 2009). Building capacity for evidence acquisition and utilization is not necessarily motivated by political ambitions to influence policy decisions directly (Blom-Hansen et al., 2021; Head, 2015; Parsons, 2004; Simon, 1955); it might also reflect a broad and continuous political agreement on policy objectives and political demand for effective policy solutions to reach such goals. Public administrations might therefore enhance their capacity to base policy decisions on evidence despite perceived barriers to research utilization (Kroll & Moynihan, 2018; Parsons, 2004). Inspired by the promises of using evidence, ministries may invest in their policy analytical capacity, understood as their capacity for: "…knowledge acquisition and utilization in policy processes" (Howlett, 2009, p. 162–63). Policy analytical capacity captures the capacity of public administrations to acquire, manage, communicate, and integrate knowledge into the decision-making stage of the policy process – something that has been emphasized as a key factor for research utilization, while lacking capacity has been associated with failure to design effective long-term policy measures (Howlett, 2009; Jennings & Hall, 2012).

In studying ministries' policy analytical capacity, we specifically focus on evidence standards, defined as the criteria used to guide decisions for accepting or rejecting available evidence as a basis for making policy decisions (Parkhurst, 2017, 161). We argue that ministries rely on evidence hierarchies to set evidence standards. We propose that ministries set evidence standards based on a consideration of the trade-off between the legitimacy derived from inclusive standards and the expected efficiency gains associated with setting exclusive evidence standards (Adam et al., 2018; Petticrew & Roberts, 2003). Ministries may set inclusive standards to gain legitimacy by signaling a willingness to include many types of knowledge in decision-making processes. Under inclusive evidence standards, many types of knowledge and organizations are likely to be included for consultation (e.g. in commissions, councils, or other types of meetings) in a broad search for the best available knowledge. As methods at the top of the evidence hierarchy, notably randomized controlled trials (RCTs), can identify average effects of policies with higher levels of certainty than studies without rigorous causal controls, ministries adhering to a rational ideal might choose to adopt exclusive evidence standards to increase research utilization and make policies more efficient, even if this excludes some types of evidence and actors not possessing such evidence from the policy process. By explicitly setting exclusive evidence standards a ministry can signal to stakeholders that it is interested in causal evidence from the top of the evidence hierarchy. In addition, exclusive evidence standards and the accumulation of causal knowledge may signal that the ministry is systematically accumulating causal evidence on policy aims that are prioritized by political decision-makers and that the ministry can credibly underpin decisions with evidence. Thus, ministries can adopt exclusive evi-

dence standards to strengthen a bridging function between evidence supply by stakeholders, in some cases by commissioning research directly, and the demand for evidence by political decision-makers and thus enhance research utilization (Howlett, 2015). We do not suggest, however, that exclusive evidence standards will always increase research utilization. Under conditions of scarce evidence and disagreement on key policy aims, the effect of exclusive evidence standards is unlikely to affect research utilization positively or at all. This is in line with other studies that emphasize conflict and a lack of evidence as barriers to research utilization (cf. Jennings & Hall, 2012). We therefore make the following proposition:

> Under conditions of available evidence and agreement on key policy aims, more exclusive evidence standards in ministries will lead to higher levels of research utilization, as they link more studies from the top of the evidence hierarchy to policy decisions.

As elaborated below, we address the proposition by studying how and why evidence standards affect research utilization in the two ministries with available evidence, policy analytical capacity, and broad political agreement on key policy aims. To address the proposition analytically, evidence standards are operationalized by studying whether ministries formally set more exclusive or inclusive criteria for using evidence in organizational documents as well as informally as captured by interviews with civil servants. As ministry practices might vary internally (different ministry divisions may have different practices and ideals for evidence use), we distinguish evidence standards set by central ministerial departments and ministerial agencies to consider organizational variation in evidence standards (Hammond, 1986). We do not consider changes in portfolio design as ministerial portfolios have been constant in the period under study for both policy domains (Sieberer et al., 2021; Fleischer et al., 2023).

In the following, the focus is specifically on evidence standards as reflected in systematic, organization-level procedures undertaken by the ministries to facilitate research utilization rather than other aspects of policy analytical capacity, such as professional analytical employment or the individual-level analytical skills of policy workers (Howlett, 2015; Migone & Howlett, 2022). Furthermore, research utilization is defined and examined as the process whereby knowledge influences political decision-making and focuses mainly on the reception and influence stages of this process, as elaborated below (Landry et al., 2003). We study differences in research utilization in the "reception stage" of research utilization (Landry et al., 2003, p. 194) both in terms of the research that ministries order and commission and the research that ministries disseminate to the public and political decision-makers. Even if crude, this distinction is important, as government ministries may apply evidence standards differently in terms of their demand and supply of evidence. Our study further aims to capture the "influence stage" of research utilization (Landry et al., 2003, p. 194) both in absolute terms by comparing the extent to which particular sources of evidence are reflected in subsequent policy decisions and in relative terms by addressing the utilization of evidence at different levels of the evidence hierarchy. Our document analysis does not capture informal discussions or uses of evidence in other phases of decision-making even if they may be relevant factors for research utilization (Knott & Wildavsky, 1980). Below, the article advances the study of evidence-based policy in two main ways. First, it develops and applies a matching method to compare levels of research utilization in policy decisions. This

addresses an analytical challenge in the evidence-based policy literature, which has focused on individual events of research utilization, often discovering barriers rather than patterns of utilization over time. Moreover, as utilization is challenging to observe directly, empirical studies have predominantly been based on perceptual data in the form of interviews or surveys with policy-makers (Boaz et al., 2009; Christensen, 2023; Head et al., 2014; Jennings & Hall, 2012; Landry et al., 2003; Newman et al., 2016; Newson et al., 2018; Toner et al., 2014). Second, the combination of content analysis and interviews with civil servants and stakeholders allows the article to examine how ministries use their policy analytical capacity and thus capture mechanisms that link evidence standards and research utilization.

Case selection

In line with international trends, the evidence movement has gained ground in Denmark, and evidence-based policy has become an ambition in many policy subsystems, including health, employment, education, and social policy (Andersen, 2020; Hansen & Rieper, 2009, 2010). Today, most Danish ministries have relatively strong capacities for knowledge acquisition and utilization, and reports have documented increasing tendencies in the ministries to consider research when preparing, enacting, and implementing policies (Arnold et al., 2015; Rambøll, 2015). Another important factor in the Danish case is that public spending levels are comparatively high. Denmark had the highest expenditure on active labor market policies as a percent of GDP among all OECD countries in 2020 (OECD Employment Database, 2023), while public expenditure on public school policy as a percentage of GDP was substantially above average in the EU27 in 2019 (Eurostat, 2023). The combination of a high capacity and high public spending levels, which incentivizes cost reduction, makes Denmark a likely context for evidence-based policy, as ministries have cost-reduction incentives and capacities to utilize evidence.

We compare evidence standards and research utilization in the Ministry of Children and Education and the Ministry of Employment from 2016 to 2021 focusing on public school policies and active labor market policies (ALMPs). The two policy areas share several background conditions for research utilization: The employment and education sectors have both been subjected to effects evaluations (Coryn et al., 2011), and evidence accumulation through systematic literature reviews have been undertaken in both sectors (e.g., through the Campbell Cooperation). Echoing the recommendation that each Danish ministry develops a knowledge strategy (DFiR, 2016), both government ministries have capacities for knowledge acquisition and utilization, affiliated research organizations, and internal agencies that work with gathering, interpreting, and applying research and data (the Danish Agency for Labour Market and Recruitment and the National Agency for Education and Quality). Furthermore, there has been a broad agreement on key policy aims in both sectors in recent decades: Due to the spread of international test regimes, student learning and well-being have been consolidated as main goals in public school policy across different governments. Concerning active labor market policy, the same goes for increasing employment through increased labor supply (Andersen, 2020, 2021; Hansen & Rieper, 2010; Ministry of Employment, 2021; STAR, 2017, 2023). On this basis, and with reference to Jennings and Hall's typology of the expected use of evidence (2012, p. 261), we regard the ministries as most-similar cases, which are expected to function as "evidence-based agencies" with

comparable and high levels of research utilization, but possible variation in outcomes due to the evidence standards adopted.

Methods

We study policy analytical capacity, evidence standards, and research utilization based on a sequential mixed methods design in which a quantitative content analysis of documents is first used to measure average levels of research utilization. The findings of the quantitative analysis are subsequently explained by drawing on interviews with ministry officials and central employment and education stakeholders (Hendren et al., 2018). Building on the definition of policy analytical capacity as knowledge acquisition and utilization in policy processes (Howlett, 2009, p. 162–63), including the extent to which public agencies have access to available, relevant, and credible evidence (cf. Jennings & Hall, 2012; Newman et al., 2016), we initially study the ministries' evidence standards including their positions toward the evidence hierarchy. Research utilization is measured both at the reception and influence stages of the policy process (Landry et al., 2003). In the reception stage, "commissioned research" captures instances where research has been ordered and funded by the ministries (LSE GV314 Group, 2014), whereas "disseminated research" captures when research has been issued on the ministries' websites. As commissioned research might also be disseminated, the two categories are not mutually exclusive. The content analysis allows us to examine research utilization as representing the "influence stage" (Landry et al., 2003), where policy-relevant information has been incorporated into policy decisions. Thereby, a "decisionistic" perspective on research utilization is applied (Christensen, 2023). In this stage, research utilization is measured as the degree to which conclusions from research publications are reflected in subsequent policy decisions. The terms "research", "evidence", and "knowledge" are used synonymously to describe written research generated systematically using both quantitative and qualitative methods (Cairney, 2016). In the following, the applied content analysis is presented, and we describe how qualitative interviews with civil servants and stakeholders were conducted and analyzed.

Content analysis and data collection

To study variation in research acquisition and utilization, a quantitative content analysis of policy decisions and research publications related to public school policies and active labor market policies in Denmark from 2016 to 2021 was conducted. The content analysis was used to assess the documents in a systematic and replicable manner based on predetermined categories and scores (Bryman, 2012; Krippendorff, 2013). The approach was designed to identify connections between research publications and policy decisions to analyze the extent and character of research utilization in the two ministries.

   An important step in this process was retrieving relevant documents for the analysis, including documents that reflect policy decisions (laws, amending acts, political agreements, orders, funds, and campaigns) and published research in the two policy areas. The latter included journal articles, research reports, evaluations, and research notes from 2015 to 2021[1], produced by Danish government agencies, consultancy firms, research organizations, think tanks, and university research units. The document retrieval exclusively focused

on documents published by actors mentioned on the ministry websites or public hearing lists from central education and employment reforms to ensure that the research publications were directly relevant to the ministries. The document retrieval, which was carried out by searching for keywords in the *Retsinformation* law database, the websites of the ministries, and the websites of relevant knowledge-producing organizations, resulted in an initial sample of 911 potentially relevant policy documents and 615 potentially relevant research publications. After screening their content, 477 documents were excluded due to irrelevancy or duplicates, while extensive policy documents, such as large policy reforms, were separated into unique policy decisions (unitized, cf. Campbell et al., 2013). This resulted in 571 policy decisions and 588 research publications retained for analysis. Table 1 displays the distribution of the documents.

Document matching analysis

The first step in the analysis was to code the policy topics, publication dates, research methods and their position in the evidence hierarchy, research providers, and whether the research publications were commissioned and disseminated by the ministries (see Appendix A for the complete list of categories). The latter two codes were used to examine differences in research utilization between the two ministries at the reception stage (cf. Landry et al., 2003). The publication dates were coded as studies have shown that timeliness is an important factor for research utilization (Nutley et al., 2007; Oliver et al., 2014), while providers were coded because policymakers' use of research from different sources tends to vary (Head et al., 2014; Jennings & Hall, 2012). Evidence standards were studied both based on strategic documents in the ministries and based on research publications' position in the evidence hierarchy, which were coded using five categories: Systematic reviews, RCTs/quasi-experimental studies, observational-analytic studies, observational-descriptive studies, and qualitative studies. The categories were inspired by the evidence hierarchy presented in Petticrew and Roberts (2003, p. 527) but were slightly modified to reflect the

Table 1 Data overview

	Active labour market policy	Public school policy
Policy decisions	325	246
Research publications	246	342
Research types		
Systematic reviews	41 (16.7%)	25 (7.3%)
RCTs/Quasi-experimental studies	43 (17.5%)	63 (18.4%)
Observational-analytic studies	29 (11.8%)	53 (15.5%)
Observational-descriptive studies	89 (36.2%)	123 (36%)
Qualitative studies	44 (17.9%)	78 (22.8%)
Total	246 (100%)	342 (100%)
Provider types		
Government agency	80 (32.5%)	20 (5.8%)
Private consultancy firm	104 (42.3%)	39 (11.4%)
Public research organization	37 (15%)	189 (55.3%)
Think tank	20 (8.1%)	20 (5.8%)
University unit	5 (2%)	74 (21.6%)
Total	246 (100%)	342 (100%)

Table 2 Distribution of disseminated and commissioned research publications[3]

Ministry			
	Ministry of Employment	Ministry of Children and Education	Total
Commissioned by the ministry			
1: Systematic reviews	11 (6.9%)	12 (6.2%)	23 (6.5%)
2: RCTs/Quasi-experimental studies	29 (18.2%)	31 (16%)	60 (17%)
3: Observational-analytic studies	23 (14.5%)	30 (15.5%)	53 (15%)
4: Observational-descriptive studies	75 (47.2%)	72 (37.1%)	147 (41.6%)
5: Qualitative studies	21 (13.2%)	49 (25.3%)	70 (19.8%)
Total	159 (100%)	194 (100%)	353 (100%)

$\chi2 = 8,913 \cdot df = 4 \cdot p = 0.063$

Disseminated by the ministry			
1: Systematic reviews	33 (19.5%)	6 (3.6%)	39 (11.7%)
2: RCTs/Quasi-experimental studies	27 (16%)	35 (21.2%)	62 (18.6%)
3: Observational-analytic studies	17 (10.1%)	29 (17.6%)	46 (13.8%)
4: Observational-descriptive studies	71 (42%)	61 (37%)	132 (39.5%)
5: Qualitative studies	21 (12.4%)	34 (20.6%)	55 (16.5%)
Total	169 (100%)	165 (100%)	334 (100%)

$\chi2 = 26,641 \cdot df = 4 \cdot p = 0,000$

evidence base in the examined policy areas. If a research publication applied more than one method, its position was based on the highest-rated method.

The conclusions from research publications were qualitatively matched with specific decisions from the policy documents. The matches were then quantified and averaged to measure the degree to which the research publications were reflected in subsequent decisions on the same topic. Inspired by previous research (cf. Knudsen, 2018; Jørgensen, 2023), the matching of research publications and policy decisions was based on three dimensions:

- Whether the research publication provides conclusions, which are subsequently addressed in a policy decision on the same topic (values 0 for "no" and 1 for "yes").
- Whether the subsequent policy decision follows or contradicts the content of the research publication (0 for contradiction, 1 for agreement).
- In case of agreement: How strong the agreement between the research publication and the decision is (1 for "moderate agreement", 2 for "strong agreement").

After coding the documents, research utilization levels were examined using two measures: "matches" and "average match values". A match was coded in cases where a policy decision addresses the same topic as a knowledge source and refers to similar target groups or contains a direct reference to a knowledge source. In terms of match value, a score of 1 indicates that the policy decision follows or contradicts the knowledge source's conclusions at a general level, that is, the direction of action recommended by the knowledge source for a similar but not necessarily identical target group. A strong agreement or contradiction is coded with the value 2 and signals that the policy decision closely follows (contradicts) the conclusions from the knowledge source by addressing the same target group and policy measure.

Knowledge sources were coded based on ten inductively identified themes for each policy domain (cf. appendix A). Match values were calculated for each theme in a knowledge source. Matches reflect the share of research publications matching one or more policy decisions (i.e., the second dimension), while average match values were calculated by dividing match values (i.e., the third dimension) by the number of observations. The measures allowed an examination of whether relevant research publications were reflected in policy decisions in the aggregate. In the process of coding, we considered introducing more matching levels, but this idea was abandoned due to challenges in coding reliability. The systematic sampling of relevant documents, the strict application of the coding scheme, and the in-depth examination of the documents, enhanced the reliability of these measures as indicators of research utilization. A team of three researchers coded the documents (see acknowledgments). To ensure the codes were applied consistently, the practices undertaken by individual coders were discussed thoroughly, and intercoder reliability scores were measured using Krippendorff's α (Krippendorff, 2013).[2] The data was analyzed using statistical software (R version 4.3.1) to cross-tabulate variables of interest and perform multiple linear regression analysis on the matches, match values, and background variables.

Interview data

The article relies on semi-structured interviews conducted in 2022 with 13 ministry officials and stakeholders related to education and employment policy to examine internal and external uses of policy analytical capacity in the ministries (cf. Appendix B). The interview respondents were purposefully selected to cover different perspectives on evidence standards and research utilization from inside and outside the ministries and to represent different organizations involved in policy processes and the production of evidence. The interviews, 45–60 minutes in length were conducted in Danish using semi-structured guides. During the interviews, key themes related to the production and utilization of evidence and recent political reforms and initiatives in the policy subsystems were addressed. The interviews were audio-recorded, transcribed, and analyzed thematically using NVivo 12 with a discussion of key themes emerging from the data (Campbell et al., 2013). Quotes from the interviews were presented to and reviewed by the interview persons for accuracy and reasonableness of interpretations. The interviews aimed to identify mechanisms and practices of research utilization in the ministries to provide explanations of the quantitative results presented below.

Comparing evidence standards and research utilization

To compare research acquisition and utilization in the Ministry of Children and Education and the Ministry of Employment, the number of research publications disseminated and commissioned by the ministries and their position in the evidence hierarchy were analyzed.

Table 2 indicates that the distribution of commissioned research publications is not significantly different between the two ministries. This is an indication that both ministries focus broadly on evaluating policies using a broad range of methods and that evidence priorities are inclusive in both ministries in terms of commissioning a broad variety of evidence. In terms of dissemination, however, there is a considerably larger share of research publications disseminated by the Ministry of Employment that apply systematic reviews compared to the Ministry of Children and Education. 19.5% of the research publications disseminated by the Ministry of Employment were systematic reviews, while the share for the Ministry of Children and Education was 3.6%. We take this as an indication that the Ministry of Employment is strategically interested in identifying systematic reviews of effect studies to facilitate evidence-based decision-making.

To examine research utilization in the ministries, we analyzed whether research publications matched specific policy decisions following the guidelines described above. Table 3 shows the resulting distribution of matches and match values across evidence hierarchy levels for the two ministries as well as general statistics about the matches.

Table 3 shows that 25% of the research publications in the dataset matched with a subsequent policy decision. The share of matched research publications is higher for the Ministry of Employment (29.3%) than for the Ministry of Children and Education (21.9%). The average match value is equally higher for the Ministry of Employment (0.63) than for the Ministry of Children and Education (0.45), indicating a higher average level of research utilization in the period. The difference in average match values is primarily characterized by a high proportion of research publications related to active labor market policy, which are in moderate agreement with subsequent decisions made by the Ministry of Employment compared to the Ministry of Children and Education.

Differences in research utilization between the two ministries were analyzed based on a linear probability model using "matches" as the dependent variable and a linear regression model using "match values" as the dependent variable.[4] Both models focus on "policy area" as the main parameter of interest. Descriptive statistics for the variables included in the analysis are provided in Table 4. Note that the total number of observations ($N=865$) is higher than the number of documents with and without a match (as displayed in Table 3) because the regression analysis includes the total number of matches and non-matches. Research publications matching more than one decision thus appear as multiple observations.

Interactions between policy area and evidence hierarchy positions were added to the regression models to examine differences between the ministries regarding the research methods they use. Both models control for potential confounders (i.e. the number of applied methods per research publication, the knowledge provider, the publication year, and whether the research publication has been disseminated and commissioned by the ministry in question). The full regression table, and corresponding logistic and ordinal logistic regressions, are shown in Appendix C–D.

Table 5 shows that average matches and match values are significantly higher ($p>0.05$) in the Ministry of Employment than in the Ministry of Children and Education in the studied

Table 3 Distribution of matches and match values across policy areas

| | Ministry | | |
	Ministry of Employment	Ministry of Children and Education	Total
Research–policy match			
No	174 (70.7%)	267 (78.1%)	441 (75%)
Yes	72 (29.3%)	75 (21.9%)	147 (25%)
Total	246 (100%)	342 (100%)	588 (100%)
$\chi2=3.727 \cdot df=1 \cdot p=0.054$			
Match value			
0: No accordance	208 (50.1%)	302 (67.1%)	510 (59%)
1: Moderate accordance	153 (36.9%)	92 (20.4%)	245 (28.3%)
2: Strong accordance	54 (13%)	56 (12.4%)	110 (12.7%)
Total	415 (100%)	450 (100%)	865 (100%)
$\chi2=31.184 \cdot df=2 \cdot p=0.000$			
Research types used by the ministry			
1: Systematic reviews	19 (26.4%)	7 (9.3%)	26 (17.7%)
2: RCTs/Quasi-experimental studies	16 (22.2%)	8 (10.7%)	24 (16.3%)
3: Observational-analytic studies	10 (13.9%)	16 (21.3%)	26 (17.7%)
4: Observational-descriptive studies	16 (22.2%)	32 (42.7%)	48 (32.7%)
5: Qualitative studies	11 (15.3%)	12 (16%)	23 (15.6%)
Total	72 (100%)	75 (100%)	147 (100%)
$\chi2=14.912 \cdot df=4 \cdot p=0.005$			
General statistics			
Total number of matches	207	148	355
Average match probability	0.499	0.329	0.410
Average match value	0.629	0.453	0.538

period. We take this as an indication that evidence standards may influence research utilization in absolute terms. To address whether research publications with a higher position in the evidence hierarchy display higher levels of research utilization, estimated marginal means were calculated for each evidence hierarchy level across the two policy areas (based on the "matches" regression model in Table 5). Figure 1 shows the marginal probabilities of a research–policy match for the five levels in the evidence hierarchy for the two ministries (exact marginal probability values are provided in Appendix E).

Table 4 Descriptive statistics for the included variables

| Observations	N	Pct.
	865	100
Dependent variables		
Research-policy match: Yes (ref. no)	355	41.0
Match value: 0	510	59.0
Match value: 1	245	28.3
Match value: 2	110	12.7
Independent variables		
Active labor market policy (ref. public school policy)	415	48.0
1: Systematic reviews	151	17.5
2: RCTs/Quasi-experimental studies	131	15.1
3: Observational-analytic studies	135	15.6
4: Observational-descriptive studies	277	32.0
5: Qualitative studies	171	19.8
Control variables		
Provider type: University unit	94	10.9
Provider type: Government agency	201	23.2
Provider type: Private consultancy firm	198	22.9
Provider type: Public research organization	320	37.0
Provider type: Think tank	52	6.0
Commissioned: Yes (ref. no)	611	70.6
Disseminated: Yes (ref. no)	542	63
Methods per publication (mean)	1.61	
Publication year: 2015	111	12.8
Publication year: 2016	155	17.9
Publication year: 2017	153	17.7
Publication year: 2018	133	15.4
Publication year: 2019	125	14.5
Publication year: 2020	85	9.8
Publication year: 2021	103	11.9

Figure 1 indicates that research publications in the Ministry of Employment that apply methods from the top of the evidence hierarchy have comparatively high probabilities of being used in a subsequent policy decision. The match probability for systematic reviews and "RCTs/Quasi-experimental studies" is 0.64 and 0.51, respectively. This echoes the results from Table 3, showing that 48.6% of the research publications used in policy decisions by the Ministry of Employment employ either systematic reviews or RCTs/Quasi-experimental studies as their highest-ranking method. In the Ministry of Children and Education, a clear relationship is not observed between higher evidence hierarchy levels and a higher probability of research utilization in the studied period. The highest match probabilities are observed in studies employing observational-analytic methods (0.42), representing 21.3% of the research publications used by the ministry.

Despite a similar policy capacity and low levels of conflict over key policy aims in both cases, the analysis shows that the Ministry of Employment exhibits higher dissemination and utilization of systematic reviews, while RCTs and quasi-experimental studies also display high utilization levels. By contrast, the Ministry of Children and Education collects and utilizes less evidence from the top of the evidence hierarchy; rather, the ministry displays a more inclusive approach to evidence.

Table 5 Regression results

Predictors	Matches		Match values	
	Estimates	SE	Estimates	SE
Policy Area: ALMP	0.176 [*]	0.072	0.270 [*]	0.117
1: Systematic reviews	0.154	0.095	0.251	0.154
2: RCTs/Quasi-experimental studies	-0.038	0.063	-0.124	0.091
3: Observational-analytic studies	0.212 [**]	0.067	0.223 [*]	0.098
4: Observational-descriptive studies	0.129 [*]	0.056	0.146	0.084
Commissioned by the ministry: Yes	0.097 [*]	0.049	0.119	0.076
Disseminated by the ministry: Yes	0.101 [*]	0.050	0.174 [*]	0.075
Interaction: ALMP and EH1	0.103	0.125	-0.105	0.198
Interaction: ALMP and EH2	0.159	0.105	0.149	0.161
Interaction: ALMP and EH3	-0.146	0.114	-0.301	0.168
Interaction: ALMP and EH4	-0.282 [**]	0.089	-0.425 [**]	0.142
Observations	865		865	
Controls	Yes		Yes	
R^2/R^2 adjusted	0.173 / 0.152		0.128 / 0.105	

Notes Linear probability model (matches) and linear regression model (match values). Both models control for publication year, the number of applied methods, provider types, and whether the ministry has commissioned and disseminated the research publication. Reference levels: Evidence hierarchy 5: Qualitative studies; Commissioned by the ministry: No; Disseminated by the ministry: No; Interaction: ALMP and Evidence hierarchy 5. $*p<0.05$, $**p<0.01$, $***p<0.001$
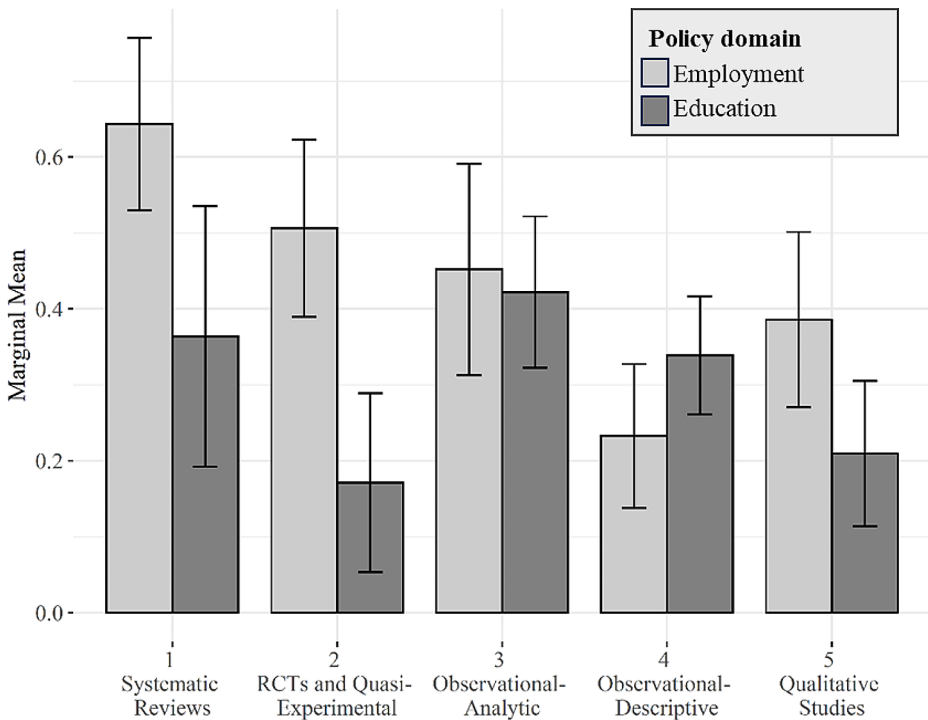
Fig. 1 Marginal match probabilities across evidence hierarchy levels

Explaining variation in research utilization

The quantitative analysis showed that the commissioning of research in the ministries is similar, but that dissemination and research utilization varies in the period with more studies from the top of the evidence hierarchy and higher levels of utilization in the Ministry of Employment compared to the Ministry of Children and Education. In this section, the analysis draws on interviews with ministry officials and stakeholders in the two policy areas to address the proposition by identifying mechanisms that link evidence standards to research utilization in the two ministries.

Policy analytical capacity and evidence standards

Interview responses from officials in the Ministry of Employment and the Ministry of Children and Education indicate that both ministries have invested in their policy analytical capacity over recent decades with a focus on collecting and interpreting evidence (Interviews 1, 2, 4, 8, 12, 13). However, there are noticeable differences in how evidence standards have been developed and applied and in how systematically evidence-based practices have developed, which can contribute to explaining the variation in research utilization in the two cases.

Historically, there have been initiatives in the Danish education sector to develop systematic evidence overviews by the Nordic Campbell Center and a What Works Clearinghouse at the Danish School of Education. However, the latter was closed in 2019, and few reviews were issued in the studied period (DPU, 2023). Instead, the Ministry of Children and Education has mainly focused on descriptive and analytic examinations of school performance data and standardized student skills testing, such as the PISA, PIRLS and National Tests (Interviews 1, 6). In the education area, interview respondents explain how the Ministry of Children and Education has been relying on inclusive evidence standards by following a broad principle of using "the best available knowledge" (Interviews 1, 2, 7). Thus, evidence is used "when relevant" rather than as a formal step in ministerial practices. A civil servant from the National Agency for Education and Quality explains that the Ministry of Children and Education continues to focus on disseminating knowledge to public schools for use in combination with local experience. Evidence in the sense of causal studies is part of the agency's daily work, but a broader palette of evidence and data is utilized, as many types of knowledge are perceived as relevant (Interview 1).

By contrast, the Ministry of Employment has pursued evidence-based policy explicitly by adopting a formal evidence strategy in 2012 – focusing on the acquisition and utilization of effect studies. The Danish Agency for Labour Market and Recruitment (STAR) has expanded its strategy to focus on investing resources in creating a knowledge bank of RCTs and quasi-experimental studies named "Jobeffekter.dk" (Andersen, 2020, 2021; STAR, 2023). Civil servants from the Ministry of Employment find that having an evidence strategy in the agency with formal evidence standards and a procedure for collecting and rating policies encourages searching for effects, closing knowledge gaps, and basing employment policies and practices on evidence at the central and local levels (Interviews 2, 8, 12). The knowledge bank reflects many aspects of the rational model of evidence-based policy by accumulating, reviewing, and rating studies to establish the impact of employment interventions on different target groups based on a hierarchy of evidence (Oliver, 2022). It includes

assessments of the quality of studies, the aggregate stock of knowledge, the dominant effect direction, and the overall level of evidence (STAR, 2023; Interview 8). The knowledge bank thus provides systematic overviews of employment effects and works as a reservoir for knowing "what works", which can underpin research utilization when civil servants advise the minister in the process preceding policy decisions. An official from the Ministry of Employment explains that the creation of the knowledge bank was motivated by a need for a stronger standing evidence-wise in political negotiations and towards the citizens. Instead of picking out random reports in political situations or when having meetings with job centers, the aim of collecting evidence systematically became a priority in the 2010s (Interview 8). The knowledge bank, which currently comprises 607 studies, can add precision to policy discussions about how interventions affect different target groups (Interviews 8, 13). Civil servants and stakeholders in the employment area explain that quantitative research with a focus on effectiveness and efficiency has had precedence over other kinds of studies – although qualitative studies also inform policy practice (Interviews 5, 10). This means that current policies generally target what the Ministry of Employment believes has the most significant effects on the labor supply for different groups of unemployed persons, based on the available evidence (Interview 8). A key aspect of the ministry's knowledge bank is that it is applied to estimate aggregate evidence levels underpinning different policy interventions, which can inform measurements of the economic costs and savings of policies (Jobeffekter, 2023). The knowledge bank thus helps filter information toward policies showing positive average effects on desired policy outcomes (Chun & Larrick, 2022; Hoefer, 2012). Particularly, evidence on the effects of business-oriented programs and caseworker interviews have influenced successive policy reforms in the 2010s, promoting a strong narrative that these interventions are effective and cost-efficient policy tools (Amilon et al., 2022; Interview 12). According to an official from the Ministry of Employment, the strong focus on quantitative studies has been driven by a desire to compare municipalities and job centers on uniform terms rather than allowing actors to base employment efforts on (random) local experiences (Interview 8).

The formal evidence strategy of the Danish Agency for Labour Market and Recruitment explicitly refers to a predetermined hierarchy of knowledge to ensure systematic and uniform conclusions in its policy recommendations (STAR 2021; 2023). In addition, efforts invested in accumulated evidence in the knowledge bank helps explain the higher share of systematic reviews disseminated and utilized by the Ministry of Employment. In addition, the institutionalization of policy analytical capacity within the Ministry of Employment to collect, interpret, and synthesize effect studies might explain the finding that the ministry does not commission many systematic reviews externally as a strong internal policy capacity makes it less relevant to order studies from external providers (cf. Table 2; Interviews 11, 13).

External aspects of policy analytical capacity

In addition to ministerial internal efforts in developing an evidence strategy with standards for research utilization and a knowledge bank, interview respondents explain that the Ministry of Employment has an ongoing external dialogue and coordination with the Ministry of Finance concerning the estimation and inclusion of derived economic effects of employment policies when estimating their costs. This economic aspect of evidence use is impor-

tant, as it produces economic incentives for evidence-based policy decisions. A ministry official emphasizes economic budgeting incentives as a driver of research utilization:

> Compared to education, we have worked so much with evidence for employment policies with the Ministry of Finance that we have co-developed cost estimation principles, which imply that if politicians make certain decisions, then sometimes they will be rewarded for following the evidence and sometimes they will experience a higher price of policy reform for not doing so. (Interview 12)

A key point here is that evidence, which indicates robust average direct or indirect effects on labor supply, is included when estimating policy costs, thus incentivizing politicians economically to utilize research and make evidence-based decisions. Because the economic effects of evidence-based policies are included in the cost estimates underpinning policy decisions, policy-makers can harvest efficiency gains from such policies regardless of whether desired policy effects are implemented and achieved in practice. While policy-makers do not always "follow the evidence", integrating policies with effects on labor supply into public budgeting provides economic incentives for doing so. As the evidence for these evidence-based policies is incorporated into budget estimates, they are moved beyond politicization in the sense that the evidence base is not normally a subject of discussion during political negotiations or contested by external stakeholders (Interview 12).

By contrast, the external aspects of research utilization in public school policy are shaped by opposing concerns. On the one hand, the Ministry of Children and Education has maintained an inclusive approach to evidence-based policy because strong professional actors and practitioners are eager to leave their mark on public school policies regardless of the evidence base, while agreement on solutions between central policy actors is regarded as important for delivering enduring policy change (Ministry of Children and Education, 2021; Interview 9). An official from the National Agency for Education and Quality explains the following:

> When we involve stakeholders, it is because they represent a teacher's perspective: What is the teachers' stance? What does it take to engage the teachers? So, we are not asking them to provide us with research advice – we get that from research institutions, or we collect it ourselves. We need them to provide their stakeholder perspective. We must have the teachers on board and ensure that the teachers see that it makes sense for them. (Interview 1)

In particular, there is a desire from the Ministry of Children and Education to mend conflicts over the 2014 Public School Reform by discussing policy problems and solutions in consultation with local actors and stakeholders in the education sector (Interview 6).

On the other hand, the central department of the Ministry of Children and Education has recently been inspired by other ministries, including the Ministry of Employment, to strengthen its focus on systematically accumulating quantitative studies and calculating the effects of education initiatives (Interviews 2, 3, 4, 5). One initiative is the platform "REFUD" (Calculation Model for Education Investments) from 2022, which enables municipalities and other actors to estimate the economic benefits of decisions in daycare and public schools based on a knowledge bank of effect studies (Jacobsen et al., 2022; REFUD,

2023). This is a step towards more exclusive evidence standards, although the main focus is currently on local government budgeting, as the Ministry of Finance does not find that the current knowledge base shows persuasive evidence that specific education policies have derived socioeconomic effects (Ministry of Finance, 2018).

In summary, the interviews indicate that research utilization is shaped both by variation in internal efforts regarding evidence standards and knowledge accumulation in "knowledge banks" and by external relations with stakeholders and dialogue with the Ministry of Finance concerning the derived economic effects of evidence-based policies. This nuances the proposition that more exclusive evidence standards will lead to higher levels of research utilization, by adding knowledge banks and economic incentives as mechanisms that affect the influence of evidence standards on research utilization in government ministries.

Discussion

The above analyses have shown that average levels of research utilization were higher for active labor market policies than for public school policies in the period 2016–2021 and attributed more exclusive evidence standards for ALMP's as well as investments in knowledge banks and economic incentives for evidence-based decisions as supplementary explanatory factors. We take this as support of our proposition in the sense that more exclusive evidence standards may increase average levels of research utilization. The influence of evidence standards, however, is contingent on ministerial policy analytical capacities to systematically link evidence to policy decisions. In particular the ability of ministries to accumulate evidence in knowledge banks appears as an important factor for accumulating evidence for the effects of early caseworker interviews with the unemployed and business interventions that have subsequently influenced successive ALMP reforms. Future research should explore whether evidence standards and knowledge banks standardize public administrations from the top-down or whether evidence continues to also flow from the bottom-up as different organizational structures convert the same evidence into different decisions and legislative activities (Hammond, 1986, 400, 403; Klüser, 2023). The qualitative analysis also identified economic incentives created by the possibility of including causal effects of policy decisions at the budgeting stage as a key mechanism in linking evidence standards to research utilization. Importantly, economic incentives attract political and administrative attention in both cases even if the Ministry of Employment has been faster and more successful in convincing the Ministry of Finance to include evidence-based effect estimates in budgeting for employment policies and reforms compared to the Ministry of Children and Education. It is possible, that variation in the importance of economic incentives stem from basic characteristic of the two policy domains. Nevertheless, exploring the dialogue between ministries and external economists further would provide a basis for determining whether particular experts have privileged access and are included more systematically than other experts in giving policy advice (cf. May et al., 2016; Migone et al., 2022). This appears particularly relevant as economic evidence and expertise has been increasingly utilized in public education policy in recent years as part of the effort to develop the REFUD database. It should be noted that the higher levels of research utilization for ALMPs identified in the above do not necessarily lead to better policy outcomes in practice. Overall expenditure has remained relatively constant in Danish job centers in the period where

evidence-based policies have been adopted, thus questioning whether evidence-based policies promote cost-efficiency in the aggregate (Amilon et al., 2022, 76). Other scholars have noted that focusing on "what works" risks coming at the expense of understanding how policies work, for whom, and under what circumstances, thereby ignoring the factors that condition policy effects and the interests of stakeholders on the ground, as well as marginalizing "non-knowledge" in the process of research utilization (Hannah et al., 2023; Head, 2008; Pawson, 2002, 2006; Parkhurst, 2017; Sanderson, 2002). Exploring the implementation and outcomes of evidence-based policies remains highly relevant even if this is beyond this study. We acknowledge that the effect of evidence standards on research utilization may be contingent on other factors than those emphasized in the above. Even if the studied policy areas display similarities in terms of policy analytical capacity, availability of evidence, and agreement on policy aims, we cannot rule out that unobserved differences between the two policy areas influence levels of research utilization. Finally, the observed agreement on broad policy aims in the two cases does not take into account that evidence-based policy in Danish public school policy has been characterized by policy conflict, in particular concerning the significant 2014 Public School Reform, which was met with opposition from many practitioners.

Conclusions

The literature on evidence-based policy contains several studies showing barriers to research utilization. However, building on the pressure for more effective and efficient policy decisions, the article has studied how and why evidence standards affect research utilization in two ministries with available evidence, policy analytical capacity, and broad political agreement on key policy aims. The article relied on existing theory emphasizing the importance of policy analytical capacity and conflict (Howlett, 2009, 2015; Jennings & Hall, 2012) to propose that more exclusive evidence standards in ministries will lead to higher levels of research utilization under conditions of available evidence and agreement on key policy aims, as they link more studies from the top of the evidence hierarchy to policy decisions.

The findings lend support to the proposition by showing that average research utilization levels were higher in active labor market policies than in public school policies in the period from 2016 to 2021 both in absolute and relative terms, and second, that evidence standards and levels of dissemination were also higher for ALMPs. By applying a matching method based on content analysis of 571 policy decisions and 588 research publications in combination with interviews, the article examined policy analytical capacity and research utilization in policy decisions in the two ministries. While 29% of available research publications matched a subsequent policy decision in the Ministry of Employment, this was the case for 22% of the research publications in the Ministry of Children and Education. Average match values, as measured on a scale from 0 to 2, were also significantly higher in employment (0.63) than in education (0.45) (cf. Table 3). Marginal match probabilities for systematic reviews and RCTs/Quasi-experimental studies were similarly higher in the Ministry of Employment than in the Ministry of Children and Education.

The qualitative part of the analysis adds nuance to the effect of evidence standards for research utilization in two respects: First, interview respondents emphasize internal aspects of analytical capacity including the importance of adopting a formal evidence strategy and the creation of knowledge banks for rating policies based on evidence. Adopting a formal evidence strategy

and systematically accumulating effect evidence in knowledge banks appear as distinguishing factors for the observed differences between the two ministries. In the Ministry of Employment, these mechanisms provided a powerful basis for identifying effective policies and utilizing research directly and indirectly to underpin the policy narrative that business programs and caseworker interviews are efficient evidence-based measures.

Second, variation in internal developments in evidence standards and policy analytical capacity interacts with external factors. In addition to policy workers' ability to acquire, manage, communicate, and integrate available knowledge into decision-making (Howlett, 2009), policy analytical capacity is also a matter of what public administrations do on an organizational level to link research to policy decisions systematically in interaction with other policy actors. The dialogue over whether and how to institutionalize evidence in public budgeting, which is normally beyond politicization in day-to-day policy-making and during policy negotiations, has developed over time and derived effects of some ALMPs are now included in economic assessments of policy costs. The inclusion of derived economic effects of some policies create economic incentives to adopt such policies, while policies for which there is no evidence for direct or derived economic effects risk being overlooked or treated as "non-knowledge" (Hannah et al., 2023). While an economic evidence-based policy agenda has been profound in Danish employment policy, education policy in Denmark is characterized by a more inclusive and politicized debate between stakeholders. Interview respondents emphasize the importance of integrating evidence-based effect measures with practical knowledge to accommodate strong professional interests and opposing ideas about what constitutes "effective" or "good" education even if the central department of the Ministry of Children and Education has recently enhanced its efforts to systematically accumulate quantitative studies. Variation and patterns in the institutionalization of evidence in public administrations, including how it is integrated into decision-making and public budgeting, should be studied in other policy sectors and countries to develop and nuance our understanding of the effect of evidence standards on research utilization in other contexts and across different government ministries.

Notes

[1] The timeframe is expanded to include 2015 for research publications to capture research possibly influencing policy decisions and debates from 2016 onwards.

[2] According to Krippendorff (2013), there is no set standard for what level of intercoder reliability is acceptable. We accepted $\alpha > 0.800$ for the background variables and $\alpha > 0.667$ for the evaluative match codes for the present study.

[3] The total number of disseminated and commissioned research publications in Table 2 is lower than the total number of research publications in the dataset because the table excludes research publications not disseminated or commissioned by the ministries.

[4] The regression analysis applies a linear probability model when focusing on the binary match variable and a linear regression model when focusing on the ordinal match value variable. Both models treat the dependent variable as continuous.

Declarations

References

Adam, C., Hurka, S., Knill, C., & Steinebach, Y. (2019). *Policy Accumulation and the democratic responsiveness trap*. Cambridge University Press.

Adam, C., Steinebach, Y., & Knill, C. (2018). Neglected challenges to evidence-based policy-making: The problem of policy accumulation. *Policy Sciences*, *51*(3), 269–290.

Amilon, A., Holt, H., Houlberg, K., Jensen, J. K., Jonsen, E. H., Kleif, H. B., Mikkelsen, C. H., & Topholm, E. H. E. (2022). *Jobcentrenes beskæftigelsesindsats*. VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd.

Andersen, N. A. (2020). The constitutive effects of evaluation systems: Lessons from the policymaking process of Danish active labour market policies. *Evaluation*, *26*(3), 257–274.

Andersen, N. A. (2021). The technocratic rationality of governance - the case of the Danish employment services. *Critical Policy Studies*, *15*(4), 425–443.

Arnold, E., Kolarz, P., Pickles, A., & Romanainen, J. (2015). *The place of Research-based evidence in policymaking*. Technopolis Group.

Baron, J. (2018). A brief history of evidence-based policy. *Annals of the American Academy of Political and Social Science*, *678*(1), 40–50.

Bhatti, Y., Hansen, H. F., & Rieper, O. (2006). *Evidensbevægelsens udvikling, organisering og arbejdsform – En kortlægningsrapport*. AKF Forlaget.

Blom-Hansen, J., Baekgaard, M., & Serritzlew, S. (2021). How bureaucrats shape political decisions: The role of policy information. *Public Administration*, *99*(4), 658–678.

Boaz, A., Fitzpatrick, S., & Shaw, B. (2009). Assessing the impact of research on policy: A literature review. *Science and Public Policy*, *36*(4), 255–270.

Boswell, C. (2009). *The political uses of Expert Knowledge: Immigration Policy and Social Research*. Cambridge University Press.

Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford University Press.

Bundi, P., & Pattyn, V. (2022). Trust, but verify? Understanding citizen attitudes toward evidence-informed policy making. *Public Administration*, (early view).

Cairney, P. (2016). *The politics of evidence-based policy making*. Palgrave Macmillan.

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, *42*(3), 294–320.

Capano, G., & Malandrino, A. (2022). Mapping the use of knowledge in policymaking: Barriers and facilitators from a subjectivist perspective (1990–2020). *Policy Sciences*, *55*(3), 399–428.

Caplan, N. (1979). The two-communities theory and knowledge utilization. *The American Behavioral Scientist*, *22*(3), 459–470.

Caplan, N., Morrison, A., & Stambaugh, R. J. (1975). *The Use of Social Science Knowledge in Policy decisions at the National Level*. Institute for Social Research, University of Michigan.

Carpenter, D., & Krause, G. A. (2015). Transactional authority and bureaucratic politics. *Journal of Public Administration Research & Theory*, *25*(1), 5–25.

Cherney, A., Head, B. W., Povey, J., Ferguson, M., & Boreham, P. (2015). Use of academic social research by public officials: Exploring preferences and constraints that impact on research use. *Evidence & Policy*, *11*(2), 169–188.

Christensen, J. (2022). When bureaucratic expertise comes under attack. *Public Administration*, (early view).

Christensen, J. (2023). Studying expert influence: A methodological agenda. *West European Politics*, *46*(3), 600–613.

Chun, J. S., & Larrick, R. P. (2022). The power of rank information. *Journal of Personality and Social Psychology*, *122*(6), 983–1003.

Coryn, C. L. S., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, *32*(2), 199–226.

Danmarks Forsknings- og Innovationspolitiske Råd (DFiR). (2016). *Anbefalinger til bedre kvalitet i politikudvikling*. Rosendahl.

Danmarks institut for Pædagogik og Uddannelse (DPU) (2023). *Dansk Clearinghouse for Uddannelsesforskning*. Retrieved July 25, 2023, from: https://dpu.au.dk/forskning/danskclearinghouseforuddannelsesforskning.

Davies, H. T. O., Nutley, S. M., & Smith, P. C. (1999). Editorial: What works? The role of evidence in public sector policy and practice. *Public Money and Management*, *19*(1), 3–5.

Daviter, F. (2015). The political use of knowledge in the policy process. *Policy Sciences*, *48*, 491–505.

Eurostat (2023). Retrieved July 25, 2023, from: https://ec.europa.eu/eurostat/databrowser/view/EDUC_UOE_FINE06__custom_4444046/default/table?lang=en.

Evans, D. (2003). Hierarchy of evidence: A framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing*, *12*(1), 77–84.

Fischer, F. (2021). *Truth and Post-truth in Public Policy*. Cambridge University Press.

Fleischer, J., Bezes, P., James, O., & Yesilkagit, K. (2023). Political time in Public bureaucracies: Explaining variation of Structural Duration in European governments. *Public Administration Review*, *83*(6), 1813–1832.

Fobé, E., & Brans, M. (2013). Policy-oriented foresight as evidence for policy making: Conditions of (mis) match. *Evidence & Policy*, *9*(4), 473–492.

Geddes, M. (2021). The webs of belief around 'evidence' in legislatures: The case of select committees in the UK House of Commons. *Public Administration*, *99*(1), 40–54.

Greany, T., & Brown, C. (2017). The evidence informed school system in England: Where should school leaders be focusing their efforts? *International Journal of Education Policy & Leadership*, *12*(3), 1–22.

Hall, J. L., & Jennings Jr., E. T. (2010). Assessing the use and weight of information and evidence in U.S. state policy decisions. *Policy and Society*, *29*, 137–147.

Hammond, T. H. (1986). Agenda Control, Organizational Structure, and Bureaucratic Politics. *American Journal of Political Science*, *30*(2), 379–420.

Hannah, A., Tchilingirian, J., Botterill, L., & Attwell, K. (2023). The role of 'non-knowledge' in crisis policy-making: A proposal and agenda for future research. *Evidence & Policy*, *19*(1), 116–130.

Hansen, H. F., & Rieper, O. (2009). The evidence movement: The development and consequences of methodologies in review practices. *Evaluation*, *15*(2), 141–163.

Hansen, H. F., & Rieper, O. (2010). The politics of evidence-based policy-making: The case of Denmark. *German Policy Studies*, *6*(2), 87–112.

Head, B. W. (2008). Three lenses of evidence-based policy. *Australian Journal of Public Administration*, *67*(1), 1–11.

Head, B. W. (2015). Toward more evidence-informed policymaking? *Public Administration Review*, *76*(3), 472–484.

Head, B. W., Ferguson, M., Cherney, A., & Boreham, P. (2014). Are policy-makers interested in social research? Exploring the sources and uses of valued information among public servants in Australia. *Policy and Society*, *33*, 89–101.

Hendren, K., Luo, Q. E., & Pandey, S. K. (2018). The state of mixed methods research in public administration and public policy. *Public Administration Review*, *78*(6), 904–916.

Hoefer, R. (2012). *Policy creation and evaluation: Understanding Welfare Reform in the United States*. Oxford University Press.

Howlett, M. (2009). Policy analytical capacity and evidence-based policy-making: Lessons from Canada. *Canadian Public Administration*, *52*(2), 153–175.

Howlett, M. (2015). Policy analytical capacity: The supply and demand for policy analysis in government. *Policy and Society*, *34*(3–4), 173–182.

Jacobsen, R. H., Andreasen, A. G., Keilow, M., Krassel, K. F., Lesner, R. V., Mortensen, N. P., Rangvid, B. S., & Verner, M. (2022). *Afdækning Af Muligheder for Udbredelse Af SØM til Indsatser på Børne- Og Undervisningsområdet*. VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd.

JenningsJr., E. T., & Hall, J. L. (2012). Evidence-based practice and the use of information in state agency decision making. *Journal of Public Administration Research & Theory*, *22*, 245–266.

Jobeffekter.dk (2023). *Jobeffektsøgning*. Retrieved July 25, 2023, from: https://www.jobeffekter.dk/jobeffektsoegning/#/.

Jørgensen, J. V. (2023). *Knowledge Utilisation Analysis: Measuring the utilisation of knowledge sources in policy decisions*. Evidence & Policy. (early view).

Klüser, J. K. (2023). From bureaucratic capacity to legislation: How ministerial resources shape governments' policy-making capabilities. *West European Politics*, *46*(2), 347–373.

Knott, J., & Wildavsky, A. (1980). If dissemination is the solution, what is the problem? *Knowledge*, *1*(4), 537–578.

Knudsen, S. B. (2018). Developing and testing a new measurement instrument for documenting instrumental knowledge utilisation: The degrees of knowledge utilization (DoKU) scale. *Evidence & Policy*, *14*(1), 63–80.

Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed.). Sage.

Kroll, A., & Moynihan, D. P. (2018). The design and practice of integrating evidence: Connecting performance management with program evaluation. *Public Administration Review*, *78*(2), 183–194.

Landry, R., Amara, N., & Lamari, M. (2001). Climbing the ladder of research utilization: Evidence from social science research. *Science Communication*, *22*(4), 396–422.

Landry, R., Lamari, M., & Amara, N. (2003). The extent and determinants of the utilization of university research in government agencies. *Public Administration Review*, *63*(2), 192–205.

Lester, J. P. (1993). The utilization of policy analysis by state agency officials. *Knowledge*, *14*, 267–290.

Lindblom, C. E., & Cohen, D. K. (1979). *Usable knowledge: Social Science and Social Problem solving*. Yale University Press.

LSE GV314 Group. (2014). Evaluation under contract: Government pressure and the production of policy research. *Public Administration*, *92*(1), 224–239.

Majone, G. (1989). *Evidence, argument, and persuasion in the policy process*. Yale University Press.

May, P. J., Koski, C., & Stramp, N. (2016). Issue expertise in policymaking. *Journal of Public Policy*, *36*(2), 195–218.

Migone, A., & Howlett, M. (2022). Assessing policy analytical capacity in contemporary governments: New measures and metrics. *Australian Journal of Public Administration*, *82*(1), 3–25.

Migone, A., McGregor, M. R., Brock, K., & Howlett, M. (2022). Super-users and hyper-experts in the provision of policy advice: Evidence from a survey of Canadian academics. *European Policy Analysis*, *8*(4), 370–393.

Ministry of Children and Education (2017). *Ny datastrategi skal få data ud at arbejde i undervisningssektoren*. Retrieved July 25, 2023, from: https://www.stukuvm.dk/aktuelt/uvm/udd/folke/2017/feb/170202%20ny%20datastrategi%20skal%20faa%20data%20ud%20at%20arbejde.

Ministry of Children and Education (2021). *Regeringen og folkeskolens parter går sammen: Folkeskolens udfordringer skal drøftes i fællesskab* Retrieved July 25, 2023, from: https://www.uvm.dk/aktuelt/nyheder/uvm/2021/maj/210528-regeringen-og-folkeskolens-parter-gaar-sammen.

Ministry of Employment (2018). *Initiativbeskrivelse: Færre proceskrav til samtaleforløb*. Retrieved July 25, 2023, from: https://bm.dk/media/7892/samlet-faerre-og-mere-enkle-proceskrav.pdf.

Ministry of Employment (2021). *Mission og vision*. Retrieved July 25, 2023, from: https://bm.dk/ministeriet/karriere-i-ministeriet/velkommen-til-beskaeftigelsesministeriet/mission-og-vision/.

Ministry of Finance. (2018). *Regneprincipper Og Modelanvendelse – Dynamiske Effekter Af Offentligt Forbrug Og Offentlige Investeringer*. Finansministeriet.

Mosley, J. E., & Gibson, K. (2017). Strategic use of evidence in state-level policymaking: Matching evidence type to legislative stage. *Policy Sciences*, *50*, 697–719.

Newman, J., Cherney, A., & Head, B. W. (2016). Do policy makers use academic research? Reexamining the two communities theory of research utilization. *Public Administration Review*, *76*(1), 24–32.

Newman, J., Cherney, A., & Head, B. W. (2017). Policy capacity and evidence-based policy in the Public Service. *Public Management Review*, *19*(2), 157–174.

Newson, R., King, L., Rychetnik, L., Milat, A., & Bauman, A. (2018). Looking both ways: A review of methods for assessing research impacts on policy and the policy utilisation of research. *Health Research Policy and Systems*, 16(1).

Nutley, S. M., Powell, A., & Davies, H. T. O. (2013). What Counts as Good Evidence? Alliance for Useful Evidence.

Nutley, S. M., Walter, I., & Davies, H. T. O. (2007). *Using evidence*. How Research Can Inform Public Services. The Policy.

OECD (2023). *Employment Database*. Retrieved July 25, 2023, from: https://stats.oecd.org/Index.aspx?QueryId=8540#.

Oliver, K. (2022). How policy appetites shape, and are shaped by evidence production and use. In P. Fafard, A. Cassola, A., E. de, & Leeuw (Eds.), *Integrating Science and Politics for Public Health* (pp. 77–101). Palgrave Studies in Public Health Policy Research.

Oliver, K., Innvar, S., Lorenc, T., Woodman, J., & Thomas, J. (2014). A systematic review of barriers to and facilitators of the use of evidence by policymakers. *BMC Health Services Research*, *14*(1), 1–12.

Parkhurst, J. O. (2016). Appeals to evidence for the resolution of wicked problems: The origins and mechanisms of evidentiary bias. *Policy Sciences*, *49*, 373–393.

Parkhurst, J. O. (2017). *The politics of evidence: From evidence-based policy to the good governance of evidence*. Routledge Studies in Governance and Public Policy.

Parsons, W. (2004). Not just steering but weaving: Relevant knowledge and the craft of building policy capacity and coherence. *Australian Journal of Public Administration*, *63*(1), 43–57.

Pawson, R. (2002). Evidence-based policy: The promise of 'realist synthesis'. *Evaluation*, *8*(3), 340–358.

Pawson, R. (2006). *Evidence-based policy: A Realist Perspective*. Sage.

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology and Community Health*, *57*, 527–529.

Powell, A., Davies, H. T. O., & Nutley, S. M. (2018). Facing the challenges of research-informed knowledge mobilization: 'Practising what we preach?'. *Public Administration*, *96*(1), 36–52.

Rambøll (2015). *Den Forskningsbaserede Viden i Politikudviklingen. Rapport til Danmarks Forsknings- og Innovationspolitiske Råd* Rambøll Management Consulting.

REFUD (2023). *Regnemodel for Uddannelsesinvesteringer* Retrieved July 25, 2023, from: https://refud.dk/.

Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public Administration*, *80*(1), 1–22.

Sanderson, I. (2011). Evidence-based policy or policy-based evidence? Reflections on Scottish experience. *Evidence & Policy*, *7*(1), 59–76.

Sieberer, U., Meyer, T. M., Bäck, H., Ceron, A., Falcó-Gimeno, A., Guinaudeau, I., Hansen, M. E., Kollveit, K., Louwerse, T., Müller, W. C., & Persson, T. (2021). The political dynamics of portfolio design in European democracies. *British Journal of Political Science*, *51*(2), 772–787.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118.

Styrelsen for Arbejdsmarked og Rekruttering (STAR) (2017). *Strategi for Styrelsen for Arbejdsmarked og Rekruttering 2017* Retrieved July 25, 2023, from: https://star.dk/om-styrelsen/publikationer/2017/01/strategi-for-styrelsen-for-arbejdsmarked-og-rekruttering-2017/.

Styrelsen for Arbejdsmarked og Rekruttering (STAR) (2018). *Agreement on a simpler employment system* Retrieved July 25, 2023, from: https://www.star.dk/en/recent-labour-market-policy-reforms/agreement-on-a-simpler-employment-system-2018/.

Styrelsen for Arbejdsmarked og Rekruttering (STAR) (2021). *STAR's Evidence Strategy*

Styrelsen for Arbejdsmarked og Rekruttering (STAR) (2023). *Evidence-based policy-making* Retrieved July 25, 2023, from: https://www.star.dk/en/evidence-based-policy-making/.

Toner, P., Lloyd, C., Thom, B., MacGregor, S., Godfrey, C., Herring, R., & Tchilingirian, J. (2014). Perceptions on the role of evidence: An English alcohol policy case study. *Evidence & Policy*, *10*(1), 93–112.

Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, *39*(5), 426–431.

Weiss, C. H. (1995). Chapter 3: The haphazard connection: Social science and public policy. *International Journal of Educational Research*, *23*(2), 137–150.

Weiss, C. H., Murphy-Graham, E., Petrosino, A., & Gandhi, A. G. (2008). The fairy godmother – and her warts: Making the dream of evidence-based policy come true. *American Journal of Evaluation*, *29*(1), 29–47.