

Ariadne-Hintergrund

Governance-Kapazitäten in der reflexiven Klimapolitik: Umfang, Rolle und institutionelle Arrangements der Evaluierung des Policy- Mixes im deutschen Gebäudesektor

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Autorinnen und Autoren



» Dr. Duncan L. Edmondson
Hertie School



» Oskar Krafft
Hertie School



» Prof. Dr. Christian Flachslan
Hertie School



» Christian van Ballegooy
Hertie School

Herausgegeben von

Kopernikus-Projekt Ariadne
Potsdam-Institut für Klimafolgen-
forschung (PIK)
Telegrafenberg A 31
14473 Potsdam

April 2024

Dieses Papier zitieren:

Duncan L. Edmondson, Oskar Krafft, Christian Flachslan, Christian van Ballegooy
(2024): Governance-Kapazitäten in der reflexiven Klimapolitik: Umfang, Rolle und insti-
tutionelle Arrangements der Evaluierung des Policy-Mixes im deutschen Gebäudesek-
tor. Kopernikus-Projekt Ariadne, Potsdam.
<https://doi.org/10.48485/pik.2023.012>

Kontakt zu den Autorinnen und Autoren:

Duncan Edmondson, duncan.edmondson@viablecities.se

Der vorliegende Ariadne-Hintergrund wurde von den oben genannten Autorinnen und
Autoren des Ariadne-Konsortiums ausgearbeitet. Die Analyse spiegelt nicht zwangsläu-
fig die Meinung des gesamten Ariadne-Konsortiums oder des Fördermittelgebers wider.
Die Inhalte der Ariadne-Publikationen werden im Projekt unabhängig vom Bundes-
ministerium für Bildung und Forschung erstellt.

Inhalt

Zusammenfassung	1
1. Einleitung	4
2. Ex-post-Evaluierungsprozesse in der reflexiven Klimagovernance	9
2.1. Governance System und institutioneller Aufbau des Evaluierungsprozesses.....	10
2.2. Der Umfang und die Qualität der Bewertungen	14
2.3. Der Einsatz von Evaluierungen im politischen Prozess	20
3. Forschungsdesign	23
4. Institutionelle Ausgestaltung der Governance der Evaluierung von Klimapolitik im Wohngebäudesektor in Deutschland	32
4.1. Verordnung und Meldepflichten.....	32
4.1.1. EU-Richtlinien und Berichterstattung.....	32
4.1.2. Sektorübergreifende Verordnung.....	34
4.1.3. Sektorale Programme	39
4.2. Governance und Ministerien	42
4.2.1. Bundesfördermittel für effiziente Gebäude (BEG).....	43
4.2.2. Gebäudeenergiegesetz (GEG)	46
4.2.3. Hilfsmittel.....	48
4.3. Berater.....	51
5. Ex-post-Evaluierung: Verfahren, Umfang, Daten und Methoden	55
5.1. Bewertungsansätze und bewertete Programme.....	55
5.1.1. Steuern/wirtschaftliche Instrumente.....	56
5.1.2. Bundesfördermittel für effiziente Gebäude (BEG).....	57
5.1.3. Gebäudeenergiegesetz (GEG)	58
5.1.4. Methodikleitfäden.....	60
5.2. Umfang der Bewertungsindikatoren.....	62
5.2.1. Wirksamkeit.....	64
5.2.2. Kosten/Kostenwirksamkeit	66
5.2.3. Verteilungseffekte	72
5.2.4. Akzeptanz	74
5.2.5. Governance.....	76
5.3. Daten	77
5.3.1. Datenqualität	77
5.3.2. Qualität der Datenberichterstattung und Zugang zu den Daten in den Bundesagenturen.....	78
5.3.3. Datenschutz und Durchsetzung von Rechtsinstrumenten.....	80
5.3.4. Einrichtung einer Gebäudedatenbank	82
5.4. Methoden.....	87
5.4.1. End- und Primärenergieeinsparungen	88
5.4.2. Auswirkungen der Anpassung.....	91
5.4.3. Interaktionseffekte	94
5.4.4. Allgemeine Empfehlungen.....	96
6. Nutzung von Evaluierungen und Auswirkungen auf den politischen Prozess	98
6.1. Analytische Beiträge zur Ex-ante-Prognose.....	98
6.2. Motivation, Beauftragung und Nutzung von Evaluierungen zur Rekalibrierung der Maßnahmen	100
6.3. Zeitplan für Evaluierungen und informelle Beratung.....	104
6.4. Verbreitung der Bewertungen und Transparenz.....	106
7. Die wichtigsten Reformoptionen	108
7.1. Institutionelle Arrangements und Koordinierung	108
7.2. Umfang der bewerteten Programme und Evaluationskriterien.....	110

7.3. Datenzugang und -qualität.....	113
7.4. Methoden	114
7.5. Nutzung der Evaluierungen und Verbreitung.....	116
7.6. Vereinbarkeit der wichtigsten Reformen mit der derzeitigen institutionellen Struktur....	118
8. Schlussfolgerungen	120
Referenzen	125
Anhang	131

Zusammenfassung

Wohngebäude sind für 11 % der Treibhausgasemissionen in Deutschland direkt verantwortlich und für bis zu 40 %, wenn Stromverbrauch mit einbezogen wird. Um die Klimaziele gemäß dem Bundes-Klimaschutzgesetz zu erreichen, ist ein höheres Ambitionsniveau der klimapolitischen Maßnahmen in diesem Sektor erforderlich. Dieser Bericht analysiert die Governance der Evaluierungsprozesse von Maßnahmen im deutschen Gebäudesektor. Wir analysieren die Rollen der beteiligten Akteure, das institutionelle Gefüge sowie die angewandten Metriken und Methoden der Evaluierungen.

Der Hintergrund stützt sich auf ein Mixed-Methods-Forschungsdesign, bei dem 14 Experteninterviews mit einer Inhaltsanalyse publizierter Evaluationsberichte und verfügbarer Literatur kombiniert werden. Der Forschungsansatz basiert zudem auf bestehender Literatur zu institutionellen sowie evaluativen Ansätzen und ermöglicht eine tiefere Einsicht in Qualität, Umfang und Methodik der ex-post Evaluation von Politikmaßnahmen im deutschen Gebäudesektor.

Wir konzentrieren uns insbesondere darauf, wie Evaluierungen die Maßnahmengestaltung und -anpassung beeinflussen. Unsere Erkenntnisse zeigen methodische, inhaltliche und institutionelle Limitierungen, welche die Erstellung und Verwendung von Evaluationen im deutschen Gebäudesektor beeinträchtigen. Wir identifizieren institutionelle und methodische Optionen zur Verbesserung dieser Prozesse mit dem Ziel, zu einer effektiveren Maßnahmengestaltung und -umsetzung beizutragen.

Die Qualität der Ex-post-Evaluierung hat direkte Auswirkungen auf die ex-ante Planung, einschließlich der Festlegung von Zielen und Strategien. Ungenauigkeiten bei der Ex-post-Evaluierung von Maßnahmen können die Präzision der Projektion der künftigen Treibhausgasminderungseffekte des geplanten Policy Mixes negativ beeinträchtigen. Gerade in Anbetracht der zunehmenden Bedeutung von Projektionsberichten in der deutschen Klimapolitik, wie sie im Entwurf der Novelle des Bundesklimagesetzes vorgesehen ist, könnten Ungenauigkeiten erhebliche Konsequenzen nach sich ziehen. Grundsätzlich sind Berichtspflichten nach dem KSG (Bundes-Klimaschutzgesetz) noch relativ neu, und die Evaluierungsprozesse sind noch nicht ausreichend angepasst worden, um alle relevanten Kriterien zu berücksichtigen und alle Maßnahmen einzubeziehen.

Der gegenwärtige Umfang der Ex-post-Evaluierungen ist begrenzt und einige zentrale Kriterien werden vernachlässigt. Besonders hervorzuheben sind hierbei bestehende Defizite in Bezug auf Verteilungseffekte, Governance-Kompetenzen und dynamische Kosteneffektivität. Zudem ist die Entwicklung aussagekräftiger Evaluierungsmethoden notwendig, um fundierte und strategische Entscheidungen in Planung und Gestaltung zu ermöglichen.

Das Gebäude-Energiegesetz (GEG) unterliegt aktuell keiner Verpflichtung zu einer Evaluation und auch zukünftig ist keine Ex-post-Evaluierung vorgesehen. Die durch Gebäudestandards entstehenden Kosten werden derzeit nicht evaluiert und erhalten somit nicht die gleiche Aufmerksamkeit wie Subventionsprogramme im Gebäudesektor. Dabei werden sowohl die potenziellen makroökonomischen Auswirkungen sowie die entstehenden Verwaltungskosten, die für eine effiziente Verwaltung und Durchsetzung von Gebäudestandards erforderlich sind, nicht berücksichtigt.

Die Bewertung von Standards steht vor mehreren zentralen Herausforderungen in Bezug auf Datenverfügbarkeit, den Vollzug und Rechenschaftspflicht. Fehlende Energieverbrauchsdaten machen es schwierig, die Auswirkungen von eingeführten Energieeffizienzstandards zu bewerten. Insgesamt fehlen Daten über die Auswirkungen von Regulierungen, was vor allem auf das Fehlen von Meldepflichten zurückzuführen ist. Ohne umfassende Daten zu Energieverbrauchsmustern wird die genaue Einschätzung der Wirksamkeit von Regulierungsmaßnahmen eine Herausforderung bleiben.

Die Wirksamkeit der Standards wird potentiell durch Mängel in der Qualitätssicherung untergraben. Selbst mit verbesserter Datenverfügbarkeit hängt die Wirksamkeit von Maßnahmen in hohem Maße von einer effektiven Arbeit der Aufsichtsbehörden ab. Die bestehenden Datenschutzgesetze erschweren der Bundesregierung den Zugang zu länderspezifischen Daten, was zu einem Mangel an Rechenschaftspflicht beiträgt. Infolgedessen wird die Effektivität der zuständigen Aufsichtsbehörden beeinträchtigt.

Die wichtigsten Empfehlungen dieses Hintergrunds sind:

- Entwicklung einer öffentlich zugänglichen anonymisierten Datenbank für den Gebäudebestand, einschließlich der aktuellen Effizienzstandards für die Gebäudehülle, des Energieträgers und der Effizienzbewertung der Wärmequelle.
- Ausbau der Ausbildung und des Fachwissens lokaler Aufsichtsbehörden und Einführung von Meldepflichten gegenüber der Bundesregierung.
- Weitere Standardisierung von Verfahren, Schätzungen und Annahmen zwischen staatlichen Auftraggebern, Forschungsinstituten und Beratungsunternehmen.
- Erweiterung der Evaluierungen um sozioökonomische Auswirkungen, dynamische Kosteneffizienz und eine explizitere Behandlung von Governance und Verwaltung.
- Erhöhung der Transparenz von Modellannahmen und -parametern bei Top-down-Evaluierungen.
- Digitalisierung von Förderprogrammdateien und verbesserte Zugänglichkeit, einschließlich des öffentlichen Zugangs.
- Verbesserung der Berichterstattung und Rechenschaftspflicht für Evaluierungen von Politikmaßnahmen, über die aktuelle Überprüfung durch den Bundesrechnungshof hinsichtlich staatlicher Haushaltsausgaben und wirtschaftlicher Effizienz hinaus.
- Einführung präziserer Methoden zur Messung der Wirksamkeit von Instrumenten zur Treibhausgasreduzierung, insbesondere von Energieeffizienzmaßnahmen.

1. Einleitung

Die Verringerung der Treibhausgasemissionen des Wohngebäudesektors ist eine große klimapolitische Herausforderung. Die Fortschritte bei der Dekarbonisierung des Gebäudebestands sind in Europa und weltweit nur langsam. Deutschland hat ehrgeizige rechtsverbindliche klimapolitische Sektorziele für die Treibhausgas-(THG)-Einsparung und hat in den letzten zehn Jahren erhebliche Fortschritte bei der Weiterentwicklung von Evaluierungsverfahren für die nationale Gebäudepolitik gemacht. Trotzdem ist die Dekarbonisierung von Gebäuden weniger weit fortgeschritten als in anderen Sektoren. Die Wirksamkeit der bisher in Deutschland umgesetzten Maßnahmen war nicht überzeugend, da die Sektorziele des Bundes-Klimaschutzgesetzes (KSG) bereits zweimal verfehlt wurden (ERK, 2022).

Für die Transformation und die Erreichung der Treibhausgasminderungsziele sind beispiellose klimapolitische Anstrengungen erforderlich. Zu den wichtigsten Reformen gehören: die Ausweitung des Policy-Mix, um auf verschiedene Markt- und Systemherausforderungen einzugehen; die Verschärfung bestehender politischer Instrumente; die Beseitigung von Markteintrittsbarrieren für neue Technologien; die Lösung von Verteilungskonflikten; die Minimierung negativer Wechselwirkungen zwischen Instrumenten; und die Bewältigung unbeabsichtigter Folgen. Entscheidend ist, dass sich die Gestaltung eines adaptiven Policy-Mix in einem Lernprozess weiterentwickeln muss, um den im Laufe der Zeit veränderten Bedingungen und Ungewissheiten gerecht zu werden (Edmondson et al., 2023, 2022). Um auf diese vielschichtigen Herausforderungen zu reagieren, sind reflexive Governance-Prozesse erforderlich: Eine zuverlässige und zeitnahe Produktion von Evaluierungsdaten ist erforderlich, um die Politik dynamisch zu aktualisieren und anzupassen; andernfalls sind die politischen Entscheidungstragenden gezwungen, sich auf eine begrenzte Evidenzbasis zu stützen, Schätzungen vorzunehmen und möglicherweise normative Ansätze zur Entscheidungsfindung anzuwenden (Fishburn, 1988).

Für eine wirksame Bewältigung der Energiewende und die Aktualisierung des Policy-Mixes sind Evaluierungsverfahren erforderlich. Da zahlreiche Instrumente zur Bewältigung dieser Herausforderungen eingesetzt werden, kann sich eine große Zahl von Instrumenten anhäufen und überlagern, was die Governance vor zusätzliche Herausforderun-

gen stellt (Berneiser et al., 2021; Meyer et al., 2021). Die Berücksichtigung der Wechselwirkungen zwischen den Instrumenten ist ein zentrales Thema bei der Bewertung von Policy-Mixes. Grundsätzlich müssen Entscheidungstragenden wissen, wie sie die Mittel – trotz begrenzter Ressourcen – am effektivsten einsetzen können, um die größtmögliche Verringerung der Treibhausgasemissionen zu erzielen.

Die Verbesserung der Kapazitäten für eine reflexive, lernende Governance kann dazu beitragen, die Wahrscheinlichkeit von Governance-Versagen zu verringern und die Anpassungsfähigkeit bei erwarteten Veränderungen der Bedingungen zu erhöhen. Zwar ist ein gewisses Maß an Ex-ante-Unsicherheit und -Komplexität unvermeidlich, wenn man das Ausmaß und die Geschwindigkeit bedenkt, mit der die Transformation des Sektors zur Erreichung der Reduktionsziele erfolgen muss (Edmondson et al., 2022). Allerdings kann eine zuverlässige Beobachtung der Auswirkungen von Maßnahmen dazu beitragen, die Zuverlässigkeit von ex-ante Projektionen zu verbessern. Dies ist insbesondere dann erforderlich, wenn bereits umgesetzte politische Instrumente auf ein noch nie dagewesenes Maß an Stringenz angehoben werden oder wenn neue Instrumente oder Gestaltungsmerkmale erprobt werden. Der Umfang der in die Evaluierungen einbezogenen Kriterien ist ebenfalls von Bedeutung für die Aussichten auf Reformen. Wenn wichtige Auswirkungen von der Evaluierung von Maßnahmen ausgeschlossen werden (z.B. Verteilungseffekte), fehlt es an zuverlässigen Belegen, auf die sich die Evaluierung und die zugrundeliegenden Annahmen über Maßnahmenoptionen stützen können. In diesen Fällen können eher diskursive oder politische Narrative eine größere Rolle spielen, auch wenn diese nicht evidenzbasiert sind. Folglich kann das, was in Bewertungen einbezogen oder ausgeschlossen wird, einen entscheidenden Einfluss auf die Gestaltung von politischen Reformen haben.

Die institutionelle Perspektive hilft, die strukturellen Dimensionen der wichtigsten Dynamiken im politischen Entscheidungsprozess zu identifizieren. Die institutionelle Perspektive bezieht sich auf die Kapazitäten, die sich auf die Gestaltung, Umsetzung, Überwachung und Bewertung der Maßnahmen auswirken. Besonders wichtig sind die Kapazitäten der Regierungsstellen und ihre Beziehungen zu den wichtigsten Akteuren, die an der Bewertung von Maßnahmen beteiligt sind (Meckling und Nahm, 2018). Diese Verfah-

rensvereinbarungen und die Koordination mit Regierungsstellen, Behörden und bevorzugt beauftragten Beratungsunternehmen bestimmen weitgehend die Qualität, den Umfang und den späteren Nutzen von Evaluierungen von Maßnahmen (Schoenefeld und Jordan, 2017). Ein besseres Verständnis dieser Zusammenhänge ist von zentraler Bedeutung, um die Rolle der Evidenz bei der Neugestaltung der Politik und den Ergebnissen der politischen Entscheidungsfindung zu entschlüsseln. Bisher gibt es nur wenige konzeptionelle Arbeiten, die untersuchen, wie institutionelle Arrangements die Produktion und Nutzung von Evidenz im klimapolitischen Prozess ermöglichen oder einschränken können (Hildén, 2011). Durch den Rückgriff auf die institutionelle Literatur trägt unser Papier dazu bei, die Lücke zwischen der klimapolitischen Literatur zum Design von politischen Instrumenten und der Rolle der öffentlichen Verwaltung zu schließen (Peters, 2012). Wir konzentrieren uns auf: (i) die formalen, strukturellen und prozeduralen Arrangements und die Koordinierung der Maßnahmenevaluierung, (ii) die Qualität und den Umfang der Evaluierungen und (iii) die Relevanz der Evaluierungen für die Entscheidungsfindung und die Maßnahmenreform.

Der Gebäudesektor ist aufgrund der Heterogenität des Gebäudebestands und der unterschiedlichen Anreize komplex. Die Heterogenität des Gebäudebestands (Wohnungstypen, Nachrüstung und Neubauten) und der Akteure (Mieter, Hauseigentümer, Vermieter, Energieversorgungsunternehmen, Installateure, Komponentenhersteller usw.) führen zu einer Vielzahl miteinander verbundener Komplexitäten bei der angestrebten Dekarbonisierung (Moore und Doyon, 2023). Eine Vielzahl an Maßnahmen und Interventionen werden eingesetzt, um unterschiedliche Möglichkeiten der Treibhausgasreduktion zu adressieren (Edmondson et al., 2020). Darüber hinaus bedeuten Verhaltensmerkmale wie Rebound-Effekte, dass die auf Gebäude ausgerichteten Maßnahmen nicht einheitlich wirken, was Annahmen erschwert und die Vorhersage der Auswirkungen schwierig macht (Galvin und Sunikka-Blank, 2016; Sunikka-Blank und Galvin, 2012). Aus der Perspektive der Energiewende ist der Gebäudesektor in der Vergangenheit weniger gut erforscht worden als andere Sektoren, wie die Stromerzeugung und der Verkehr (Köhler et al., 2019). Neuere Beiträge haben begonnen, diese Lücke zu schließen (Edmondson et al., 2020; Moore und Doyon, 2023)

Die sektoralen institutionellen Perspektiven sind noch nicht ausreichend erforscht, insbesondere im Gebäudesektor. Die meisten der vorhandenen Beiträge in der institutionellen Literatur konzentrieren sich auf die „Klimapolitik“ im weiteren Sinne und untersuchen die Unterschiede zwischen den nationalen Klimainstitutionen (Dubash et al., 2021; Finnegan, 2022; Guy et al., 2023). Folglich wurde den institutionellen Konfigurationen und Kapazitäten für klimapolitische Entscheidungsprozesse auf sektoraler Ebene bisher wenig Aufmerksamkeit geschenkt. Einige Beiträge in der Literatur zu Nachhaltigkeits-Transformationen mit sektoralem Fokus haben versucht, die Rolle von Institutionen besser zu integrieren, aber diese werden oft als kontextuelle Faktoren und nicht als expliziter Fokus des politischen Prozesses betrachtet (Brown et al., 2013; Gillard, 2016; Kern, 2011). Insbesondere gibt es nur wenige Forschungsarbeiten, die sich mit den institutionellen Konfigurationen, die politische Prozesse auf sektoraler Ebene strukturieren, befassen. Um diese Lücke zu schließen, kartieren wir die Institutionen, die Maßnahmen im deutschen Wohngebäudesektor gestalten und umsetzen, und legen dabei besonderes Augenmerk auf die Rolle der Ex-post-Evaluierung der Maßnahmen in diesen Prozessen.

Der Bericht stützt sich auf ein Mixed-Methods-Forschungsdesign und kombiniert mehrere Quellen von Primärdaten. Wir kombinieren eine institutionelle Bestandsaufnahme, eine Inhaltsanalyse veröffentlichter Berichte und Expertenbefragungen, um die Entwicklung und den Einfluss institutioneller Vereinbarungen auf die Evaluierungs- und Reformprozesse von Maßnahmen zu analysieren. Wir stellen unseren Bewertungsansatz in Abschnitt 2 vor, während unser Forschungsdesign in Abschnitt 3 detailliert beschrieben wird. In Abschnitt 4 wird die institutionelle Ausgestaltung der Wohnungsbaupolitik skizziert. In Abschnitt 5 werden dann Inhaltsanalysen und Interviews kombiniert, um die derzeitigen Evaluierungsprozesse in Bezug auf Umfang und Qualität zu bewerten. Abschnitt 6 befasst sich mit der Nutzung und Verbreitung von Evaluierungsverfahren in politischen Entscheidungsprozessen. In Abschnitt 7 werden wichtige Empfehlungen für Reformen vorgeschlagen, bevor in Abschnitt 8 Schlussfolgerungen gezogen werden.

Wir erörtern die Auswirkungen auf die Gestaltung von politischen Maßnahmen im deutschen Wohngebäudesektor und geben methodische und institutionelle Empfehlungen für die Erreichung der klimapolitischen Ziele bis 2030 und darüber hinaus. Der

Bericht geht der Frage nach, wie institutionelle Konfigurationen die Erstellung zuverlässiger Ex-post-Evaluierungen und die Nutzung von Evidenz in dynamischen und transformativen Maßnahmengestaltungsprozessen erleichtern oder einschränken können. Deutschland verfügt über gut etablierte und transparente institutionelle Arrangements in diesem Bereich, was es zu einem empirisch ergiebigen Fall macht. Damit trägt unsere Forschung zu der wichtigen Diskussion über den Bedarf an besserer Evidenz im deutschen Bausektor bei (Singhal et al., 2022), mit einer umfassenden Untersuchung der institutionellen Konfiguration, die die Evaluierung strukturiert, und mit Schlüsselempfehlungen zur Verbesserung der Qualität und Nutzung von Evidenz.

2. Ex-post-Evaluierungsprozesse in der reflexiven Klimagovernance

In diesem Abschnitt wird unsere Vorgehensweise für eine systematische Analyse der Evaluierungsprozesse von Klimaschutzmaßnahmen im deutschen Wohnungssektor beschrieben. Wir führen eine analytische Heuristik ein, um die Rolle der Evaluation in politischen Prozessen zu untersuchen. Die Heuristik entwickelt drei Rahmen, die unsere Untersuchung während der gesamten Forschung leiten. Erstens, eine Kategorisierung und Nutzung von Institutionen, um die Konfiguration von Governance und Evaluation abzubilden. Zweitens: Herausforderungen bei der Maßnahmengestaltung und methodische Überlegungen zur Bewertung des Umfangs und der Qualität der derzeitigen Evaluierungsverfahren. Und schließlich die Schlüsselfaktoren für den Einsatz von Evaluierungen im Politikprozess.

Eine wirksame Überwachung, Bewertung und Anpassung der Maßnahmen ist für eine vorausschauende Gestaltung des Policy-Mixes und eine reflexive Governance erforderlich. Eine erfolgreiche Neukalibrierung der Maßnahmengestaltung im Laufe der Zeit ist erforderlich, um sich an veränderte Bedingungen anzupassen und aus der bisherigen Umsetzung zu lernen (Morrison, 2022). In ähnlicher Weise ist eine Überarbeitung der Klimapolitik angesichts der inhärenten Unsicherheiten in den Treibhausgasbudgets erforderlich (Michaelowa et al., 2018). Aus dem jüngsten Bericht des IPCC AR6 geht hervor, dass das verbleibende Treibhausgasbudget für das 1,5-Ziel nur 50 % des zuvor erwarteten Ziels beträgt (IPCC 2023). Der jüngste Bericht des Europäischen Wissenschaftlichen Beirats zum Klimawandel empfiehlt EU-weite Klimaziele für 2040 (ESABCC, 2023), was eine weitere Änderung der deutschen KSG-Ziele erforderlich machen könnte. Dies unterstreicht die Notwendigkeit einer Erhöhung des politischen Ehrgeizes und der Rigorosität von Maßnahmen.

Die Rekalibrierung von Maßnahmen wird durch die zuverlässige und rechtzeitige Generierung von Evidenz unterstützt. Ohne Evidenz werden Entscheidungen naturgemäß zu annahmegestützten Schätzungen auf der Grundlage begrenzter Daten. Zwar sind der Bereitstellung von Daten gewisse inhärente und unvermeidliche Grenzen gesetzt, insbesondere in einem sich rasch verändernden Umfeld, doch sollten selbst Schätzungen auf

den besten verfügbaren Daten beruhen. Evidenz kann mehrere Funktionen der reflexiven Governance verbessern: (i) Überwachung und Anpassung der Stringenz von Instrumenten; (ii) Überwachung der Einhaltung/Umgehungseffekte; und (iii) Verringerung negativer Wechselwirkungen und Überlagerung komplexer (und potenziell widersprüchlicher) Instrumentenmixe. Obwohl die Bedeutung von Maßnahmenevaluierungen in der breiteren wissenschaftlichen Literatur zunehmend anerkannt wird (Fujiwara et al., 2019), gibt es nur wenige Studien, welche die Auswirkungen und Ergebnisse von Evaluierungen der Klimapolitik in der Praxis systematisch zusammenstellen und bewerten. Einige Studien haben systematische Übersichten über Ex-post-Evaluierungen der Klimapolitik durchgeführt (Auld et al., 2014; Fujiwara et al., 2019; Haug et al., 2010; Huitema et al., 2011). Diese Studien konzentrieren sich jedoch in erster Linie auf die Evaluierungsergebnisse, während die Qualität der staatlich durchgeführten Evaluierungen weitgehend unberücksichtigt bleibt.

Dieser Bericht untersucht den institutionellen Aufbau, die Qualität und die Nutzung von Maßnahmenevaluierungen. Wir tun dies, indem wir uns auf drei analytische Rahmen konzentrieren, die sich auf Evaluierungsprozesse beziehen: (i) der Governance-Rahmen und die institutionelle Konfiguration von Evaluierungsprozessen im deutschen Wohnungsbausektor; (ii) der Umfang und die Qualität von öffentlich zugänglichen Evaluierungen, (iii) die Nutzung von Evaluierungen im Politikprozess.

2.1. Governance System und institutioneller Aufbau des Evaluierungsprozesses

Wir kategorisieren die institutionelle Konfiguration der sektoralen Governance in formale und strukturelle Elemente. Zu den formellen Institutionen gehören wichtige Verordnungen (Gesetze und Rechtsakte), Strategien und Vorschriften (Kaufmann et al., 2018). Die Einführung wichtiger Verordnungen erfordert häufig die Aufstellung von Programmen und politischen Instrumenten, um die verankerten Ziele zu erreichen. Zu den strukturellen Institutionen gehören Ministerien und Bundesbehörden, die mit der Gestaltung und Umsetzung beauftragt sind (Thelen, 1999) und unterstützende Institutionen, die bei der Umsetzung und Evaluierung der politischen Instrumente helfen (Edmondson 2023). Strukturelle Institutionen sind auf die Erreichung der politischen Ziele ausgerichtet, die in formellen Institutionen verankert sind (Steinmo und Thelen, 1992). Dies kann

eine Neukalibrierung der staatlichen Kapazitäten und Zuständigkeiten innerhalb der bestehenden Struktur beinhalten (Hacker et al., 2015), oder die Bildung neuer Vereinbarungen, die auf die Erreichung eines bestimmten Ziels (oder einer bestimmten Funktion, z.B. einer Überwachungskommission) ausgerichtet sind. Diese institutionellen Elemente interagieren durch Verfahrensregeln und -praktiken (Skogstad, 2023), wie zum Beispiel Delegation (Kuzemko, 2016; Tosun et al., 2019).

Evaluierungsprozesse spielen eine Schlüsselrolle in der institutionellen Konfiguration der Klimapolitik. Institutionen formen und werden durch die Maßnahmengestaltung geformt (Peters 2000). Maßnahmengestaltung ist ein sich ständig weiterentwickelnder, nichtlinearer Prozess (Edmondson et al. 2019), der sich innerhalb einer institutionellen Konfiguration abspielt (Howlett und Ramesh, 2003). Wir betrachten drei verschiedene Phasen der Maßnahmengestaltung: (i) Agenda-Setting; (ii) Maßnahmenformulierung und -gestaltung und (iii) Maßnahmenumsetzung. Neuere Forschungen von Guy et al. (2023) zeigen, dass nationale Klimainstitutionen die verschiedenen Phasen des politischen Prozesses beeinflussen, doch es fehlt den betrachteten „Klimainstitutionen“ eine klare Definition hinsichtlich ihrer strukturellen, formalen oder prozeduralen Funktion. Die Bewertung, Überwachung, Reform und Einbeziehung von Erkenntnissen in die Entscheidungsfindung spielen in jeder Phase eine Schlüsselrolle. Die explizite Rolle der Evaluierung innerhalb formeller und struktureller institutioneller Arrangements ist bisher jedoch unerforscht geblieben. Wir konzentrieren uns auf diese Lücke in der aktuellen Literatur und befassen uns eingehend mit den Phasen der Maßnahmengestaltung und -umsetzung sowie mit der Rolle, die Evaluierungen bei der Gestaltung von Reformen spielen können.

Beim Agenda-Setting geht es darum, wie die Klimapolitik vom Staat als politisches Problem verstanden wird. Diese Phase wird weitgehend von politischen Akteuren dominiert und die Interessen konkurrierender politischer Parteien sowie Interessengruppen spielen in den Entscheidungsprozessen eine zentrale Rolle (Howlett, 1998). In diesen Prozessen sind in der Regel weniger Bürokraten, technische Experten oder sektorale Fachleute vertreten, was den Prozess der Aktualisierung von Programmen bei geplanten Überarbeitungsschritten politisieren kann (Lockwood et al., 2017). Akteure können diese Prozesse beeinflussen, indem sie die Sichtbarkeit und Bedeutung von Themen und die Formulierung von Interessen durch den öffentlichen Diskurs oder die Medien erhöhen

(Tversky und Kahneman, 1981). Für den Zweck dieses Berichts gehen wir nicht auf diese Dynamik ein. Wir konzentrieren uns auf die Rolle, die Evaluierung und Berichterstattung bei der Beeinflussung dieser Phase des politischen Prozesses haben können, ohne kausale Zusammenhänge zwischen Evaluierungsergebnissen und Auswirkungen auf die Agenda zu identifizieren.

Bei der Politikformulierung kommen politische Akteure, Bürokraten, öffentliche Einrichtungen und Interessengruppen zusammen. Ein politisches Teilsystem wird durch ein inhaltliches Thema und einen geografischen Geltungsbereich definiert und setzt sich aus einer Reihe von Akteuren zusammen, darunter Beamte aller Regierungsebenen, Vertretende verschiedener Interessengruppen und Forschende (Howlett et al., 1996; Sabatier und Weible, 2014). Das politische Subsystem im Kontext des Wohnungssektors in Deutschland besteht aus den verschiedenen Stakeholdern, die mit dem Ziel der Dekarbonisierung von Gebäuden befasst sind. Es umfasst Beamte verschiedener Regierungsebenen, Vertretende von Interessengruppen und Forschende (Mukherjee et al., 2021). Während Ministerien eine wesentliche Rolle bei der Überwachung wichtiger Verordnungen spielen, sind sie nicht auf spezifische Mandate beschränkt, da andere politische Ziele um Unterstützung und Ressourcen konkurrieren (Öberg et al., 2015). Sektorspezifische politische Programme werden in der Regel unter der Verantwortung von Ministerien entworfen und abgestimmt, aber auch Beiträge von Beratungsausschüssen, Forschungsinstituten, Beratungsfirmen und Think Tanks beeinflussen die Abstimmung und Aktualisierung von Instrumenten und Programmen im Laufe der Zeit (von Lüpke et al., 2022). Verlässliche Informationen sind für eine wirksame Rekalibrierung und Aktualisierung politischer Instrumente und Programme von entscheidender Bedeutung und erfordern umfassende Bewertungen, die ein breites Spektrum von Indikatoren abdecken, um Entscheidungen über die Gestaltung und Rekalibrierung zu treffen.

Die Umsetzungsphase spielt eine entscheidende Rolle bei der Bestimmung des Tempos und der Richtung des sozio-technischen Wandels, der sich aus der Formulierung und Gestaltung der Maßnahmen ergibt. In der Literatur über Maßnahmengestaltung und Institutionen wird die Bedeutung der Umsetzung über die effektive Gestaltung der politischen Elemente hinaus oft übersehen. In der Praxis werden die beabsichtigten Auswirkungen politischer Maßnahmen oft nicht erreicht, weil institutionelle Faktoren die

Auswirkungen auf die Zielgruppe beeinflussen (oder begrenzen) (Patashnik, 2009). Während die Umsetzung in jüngster Zeit in Bezug auf die Unterschiede zwischen den nationalen Klimainstitutionen betrachtet wurde (Guy et al., 2023), entspricht dies nur einer breiten Konzeptualisierung der „staatlichen Kapazitäten“. Genauer gesagt sind Verwaltungsressourcen und die Fähigkeit des effektiven Vollzugs sind entscheidende Kapazitäten, die die Auswirkungen der Maßnahmen beeinflussen (Edmondson 2023). Diese Aufgaben werden in der Regel von delegierten Akteuren wie Bundesbehörden, öffentlichen Einrichtungen, dezentralen Verwaltungen oder kommunalen Behörden wahrgenommen (Cairney et al., 2016; Hendriks und Tops, 2003). Die Delegation ist manchmal für eine wirksame Implementierung erforderlich, insbesondere wenn politische Maßnahmen eine lokale Verwaltung erfordern (Jordan et al., 2018). Idealerweise sollten die delegierten Akteure von zentralisierten Bundesbehörden koordiniert oder beaufsichtigt werden, um die Rechenschaftspflicht und die Durchsetzung zu verbessern und die Umgehung von Vorschriften zu verringern. Dementsprechend ist die Evaluierung der Governance-Kapazitäten für eine wirksame Maßnahmensumsetzung von wesentlicher Bedeutung. Die für die Maßnahmenevaluierung zuständigen Institutionen sollten effizient, effektiv und gut koordiniert sein. Daher sollten bei der Evaluierung von Instrumenten und Programmen explizit die Governance-Kapazitäten für eine wirksame Umsetzung von Maßnahmen berücksichtigt werden.

Evaluierungen können in Bezug auf den Umfang der gemessenen Indikatoren und die Methodik eine einseitige Auswahl vornehmen. Da Evaluierungen erhebliche Probleme von Maßnahmen und ihrer Umsetzung aufdecken können, besteht die Gefahr voreingenommener oder nur selektiv durchgeführter Evaluierungen (Bovens et al., 2008; Mastenbroek et al., 2016; Schoenefeld und Jordan, 2017). So ist es zum Beispiel weniger wahrscheinlich, dass Regierungsstellen mit einer bestimmten politischen Agenda die etablierten politischen Ziele während der Evaluierung in Frage stellen (Huitema et al., 2011). Dies kann die Generierung von Erkenntnissen einschränken, die andernfalls zu Reformen motivieren würden. Unabhängige Evaluierungen können wiederum durch einen eingeschränkten Datenzugang (z.B. durch Datenschutzgesetze) eingeschränkt sein, und selbst wenn sie veröffentlicht werden, haben sie möglicherweise weniger Einfluss auf den politischen Prozess als offiziell durchgeführte oder in Auftrag gegebene Evaluierungen (Hildén, 2011).

2.2. Der Umfang und die Qualität der Bewertungen

Nachdem wir den institutionellen Aufbau von Evaluierungsprozessen skizziert haben, gehen wir nun auf die Indikatoren ein, die in umfassenderen Evaluierungen berücksichtigt werden sollten, sowie auf Faktoren, die die Qualität von Evaluationen beeinflussen.

Umfang

Evaluierungen sollten ein breiteres Set an Bewertungskriterien berücksichtigen, damit genügend Daten für wirksame Politikreformen zur Verfügung stehen. Um die Reformen wirksam zu steuern, müssen politische Entscheidungsträger über ausreichende Daten verfügen (Edmondson et al., 2023). Planungsprozesse beruhen in der Regel auf Schätzungen, begrenzten Daten, Modellierungen und Annahmen, die die Komplexität der realen sozioökonomischen Bedingungen nicht erfassen. Bei der Evaluierung der umgesetzten Maßnahmen sollten die berücksichtigten Auswirkungen über die Energieeinsparungen und die Kosten des Programms hinausgehen, wie etwa die politische Machbarkeit, Wohlfahrts- und Verteilungseffekte (Zachmann et al., 2018). Dafür sollte ein breites Spektrum an empirischen Daten erhoben werden, das auch potenzielle Hindernisse für die Umsetzung, politische Akzeptanz und Governance-Herausforderungen umfasst. Wir stützen uns auf die von Edmondson et al. (2022) entwickelten Bewertungskriterien als Rahmen für den Umfang von Maßnahmenevaluierungen für den Wohngebäudesektor (Tabelle 1).

Herausforderung	Komponenten	Analytische Elemente für den Wohnungsbausektor
Effektivität	<i>Energieeinsatz</i>	<ul style="list-style-type: none"> - Energieverbrauch vor der Umsetzung. - Energieverbrauch nach der Umsetzung. - Rebound-Effekte.
	<i>THG-Einsparung</i>	<ul style="list-style-type: none"> - Treibhausgasintensität des Energieträgers (Emissionsfaktor). - Veränderung im Laufe der Zeit (Zusammensetzung des Strom-/Energimixes).
	<i>Interaktions-effekte</i>	<ul style="list-style-type: none"> - Positive Wechselwirkungen - Synergieeffekte zwischen den Instrumenten. - Negative Wechselwirkungen - Konflikte, die die Energieeinsparungen verringern. - Erforderliche Wechselwirkungen - Bedingungen/Eingriffe, die erforderlich sind, damit andere Maßnahmen die angenommenen Wirkungen erzielen (z.B. Mindesteffizienzstandards für den Betrieb von Wärmepumpen).
Kosteneffizienz	<i>Dynamische Kosteneffizienz</i>	<ul style="list-style-type: none"> - Marktversagen, u.a.: Kurzsichtigkeit der Verbraucher, Spillover-Effekte durch „Learning by Doing“, Spillover-Effekte in Forschung und Entwicklung, externe Effekte des Netzes. - Systemisches Versagen: Koordinations-, Strategie- und Lieferkettenversagen. - Auswirkungen auf die Investitionen. - Kosteneffizienz im Laufe der Zeit unter sich ändernden Bedingungen (d.h. Zusammensetzung des Energimixes). - Makroökonomische Auswirkungen im Zeitverlauf. - Zusätzlichkeit des Programms/Instruments.
	<i>Statische Kosteneffizienz</i>	<ul style="list-style-type: none"> - Marginale Vermeidungskosten.
Steuerliche Belastung	<i>Kosten/Einnahmen für den Staat</i>	<ul style="list-style-type: none"> - Steuerliche Kosten/Einnahmen, die durch die Maßnahmen/das Programm entstehen.
Verteilung	<i>Bevölkerung</i>	<ul style="list-style-type: none"> - Verteilung von Kosten und Nutzen auf die Bevölkerung. - Ausrichtung der Subventionen. - Direkte Verteilungseffekte (Subvention). - Indirekte Verteilungseffekte (Energieverbrauch). - Marktpreis der Energie. - Umlage der Kosten auf den Vermieter/Mieter.
	<i>Firmen</i>	<ul style="list-style-type: none"> - Verteilung der Kosten auf die Unternehmen und Auswirkungen auf die nationale Wettbewerbsfähigkeit. - Schaffung von Arbeitsplätzen.
Akzeptanz	<i>Bevölkerung</i>	<ul style="list-style-type: none"> - Akzeptanz in den Bevölkerungsgruppen.
	<i>Firmen</i>	<ul style="list-style-type: none"> - Akzeptanz bei den Interessengruppen/Stakeholdern der Branche.
	<i>Politisch</i>	<ul style="list-style-type: none"> - Unterstützung durch die Regierungsparteien. - Kohärenz zwischen der föderalen Regierung und den dezentralen Behörden.
Governance	<i>Administrative/informationelle Anforderungen</i>	<ul style="list-style-type: none"> - Überwachungs- und Durchsetzungskapazitäten. - Einhaltungquoten. - Bereitstellung von Informationen und Daten.

Tabelle 1: Herausforderungen bei der Gestaltung des Policy-Mix und entsprechende analytische Komponenten für die Bewertung. Angepasst übernommen von Edmondson et al. (2023)

Wirksamkeit insbesondere bei der THG-Reduktion ist die grundlegendste Kategorie bei der Maßnahmenevaluierung. Wie die Wirksamkeit in verschiedenen Dimensionen gemessen wird, kann anhand absoluter Kennzahlen (wie der insgesamt eingesparten Energie oder der THG-Einsparung) oder anhand von Ersatzindikatoren (wie der Anzahl der installierten Heizungsanlagen) als Prozentsatz des Gesamtbestands dargestellt werden (Edmondson et al., 2023). Die erste Kategorie liefert einen besser quantifizierbaren Wert für die Eingabe in Wirkungsmodelle und Bewertungsinstrumente, aber die Berechnungsmethodik zur Schätzung dieser Werte ist sehr wichtig. So können beispielsweise Ungenauigkeiten beim Ausgangswert des Energieverbrauchs vor der Umsetzung politischer Maßnahmen und Maßnahmen zur Treibhausgasreduzierung zu einer Überschätzung der Auswirkungen führen, d.h. zu einem „Prebound-Effekt“ (Galvin und Sunikka-Blank, 2016; Sunikka-Blank und Galvin, 2012).

Die Kosteneffizienz ist ein weiteres zentrales Kriterium. Politische Maßnahmen verursachen Kosten, die entweder in die Steuerhaushalte fließen oder sozialisiert und auf die gesamte Wirtschaft und Bevölkerung verteilt werden. Die Kosteneffizienz von Maßnahmenoptionen war ein Hauptschwerpunkt der deutschen Evaluierungen in der Vergangenheit (Rosenow und Galvin, 2013). Die Art und Weise, wie diese Kosten berechnet werden, wirkt sich erheblich auf die Rentabilität einiger Maßnahmen gegenüber anderen aus. Regulierungsmaßnahmen haben zwar keine direkten Kosten, verursachen aber Verwaltungskosten (Baek und Park, 2012). Die Einhaltung der Vorschriften verursacht auch kurzfristige Kosten für Unternehmen und/oder Haushalte, aber dieses sogenannten „versunkenen Kosten“ amortisieren sich über einen längeren Zeitraum durch die Senkung der Heizkosten (Qian et al., 2016). Dies schafft gespaltene Anreize zwischen Vermietern und Mietern, je nachdem, wer die Kosten der Einhaltung der Maßnahmen trägt und wer davon profitiert (Melvin, 2018). Marktbasierte Maßnahmen verursachen oft keine direkten Kosten für den Staatshaushalt, aber sie verursachen Kosten für verschiedene Bevölkerungsgruppen (Wang et al., 2016). Um ein breites Spektrum an politischen Instrumenten effektiv zu bewerten und Optionen sinnvoll zu vergleichen, ist daher die Evaluierung eines breiteren Spektrums von Faktoren erforderlich, darunter makroökonomische Kosten, Verteilungseffekte und Governance-Anforderungen (Verwaltungs-/Durchsetzungskosten).

Evaluierungen sollten ausdrücklich dynamische Effekte und gesellschaftliche Auswirkungen berücksichtigen. Die Betrachtung der Kosteneffizienz aus einer dynamischeren Perspektive erfordert die Berücksichtigung von Innovationen, Markt- und Systemversagen sowie der Wechselwirkungen zwischen politischen Instrumenten im Laufe der Zeit. Beispielsweise ist die Treibhausgasreduzierung bei sehr hohen Sanierungsstandards für einen Haushalt, der noch einen gasbasierten Energieträger für die Raumheizung verwendet, am größten. Mit der Dekarbonisierung des Energieträgers sinkt jedoch die THG-Einsparung durch die energetische Sanierung. Da die höchsten Sanierungsraten bereits hohe Grenzvermeidungskosten haben (Galvin, 2023), würden diese dynamisch betrachtet im Laufe der Zeit noch höher werden. Ähnlich verhält es sich mit Markt- und Systemversagen (Weber und Rohracher, 2012), welche berücksichtigt werden sollten, um potenzielle Probleme und Engpässe zu bewerten, die die Umsetzung oder Kosteneffizienz behindern würden (Edmondson et al., 2023). Diese Auswirkungen müssen antizipiert und (so weit wie möglich) durch die Gestaltung der Maßnahmen abgemildert werden. Sie sollten jedoch auch einen zentralen Aspekt der Maßnahmenevaluierung darstellen. Die umfassenderen Auswirkungen der Maßnahmengestaltung haben gesellschaftliche Implikationen. So kann beispielsweise die Ausrichtung von Subventionen zu einer regressiven Verteilung führen, wenn die Mittel nur Hausbesitzenden zur Verfügung gestellt werden. Ebenso wichtig ist es, Anreize für Vermietende zu schaffen, die Effizienz und den Energieträger von Wohnungen zu verbessern, ohne dass den Mietenden dadurch negative Kosten entstehen (George et al., 2023).

Die Governance-Anforderungen und die Fähigkeiten des Staates zur Umsetzung und Verwaltung der Klimapolitik sind wichtige Aspekte, um Wirksamkeit zu gewährleisten.

Die Evaluierung sollte sich nicht nur auf die geschätzten Auswirkungen des Instruments selbst konzentrieren und allgemeine Annahmen über die Umsetzung und Verwaltung treffen. Stattdessen sollten die Evaluierungen auch die Kapazitäten des Staates zur wirksamen Umsetzung und Durchsetzung der Maßnahme einbeziehen. Einige Maßnahmen hängen entscheidend von der Qualität des Monitorings und Vollzugs ab, um sicherzustellen, dass sie die erwarteten Energieeinsparungen und Treibhausgasreduzierungen erzielen. Insbesondere müssen Regulierungsmaßnahmen glaubwürdig durchgesetzt werden, da andernfalls Nichteinhaltung und Umgehung wahrscheinlich sind (Garmston und Pan,

2013; Hovi et al., 2012; Lu et al., 2022). Eine wirksame Umsetzung erfordert ausreichende Kapazitäten (technische Fähigkeiten und Ausbildung, Mittelzuweisung, Beibehaltung/Wechsel des Personals), Koordinierung zwischen Ministerien und Behörden (horizontal) und zwischen verschiedenen Regierungsebenen (vertikal). Evaluierungen sollten versuchen, diese Aspekte der politischen Koordination und Governance-Kapazitäten ausdrücklich zu bewerten (Edmondson 2023).

Qualität

Die Qualität der Evaluierungen hängt in hohem Maße von der Verfügbarkeit und Qualität der Eingangsdaten ab. Unzuverlässige Inputdaten machen Schätzungen und Näherungswerte für die Auswirkungen von Instrumenten und Programmen erforderlich. Dies kann dazu führen, dass die Wirksamkeit des aktuellen Policy-Mix überschätzt wird (van den Bergh et al., 2021). In einigen Fällen wird ein schlecht funktionierender Policy-Mix möglicherweise nicht als solcher erkannt. Dementsprechend können ineffektive Instrumentendesigns und Programme fortbestehen und nur kleine inkrementelle Änderungen erfahren (Jacobs und Weaver, 2015). Genauere Erkenntnisse, die unerwünschte Auswirkungen oder die Kosten(in)effizienz des bestehenden Programms aufzeigen, können dagegen Anlass für radikalere Reformen sein. Dementsprechend sollten Evaluierungsprozesse die Qualität und Zuverlässigkeit der Eingangsdaten ermitteln und entsprechende Mängel in der Datenlage in den Evaluierungsergebnissen angegeben werden. Noch wichtiger ist, dass festgestellte Probleme bei der Datenverfügbarkeit angegangen werden sollten, um zuverlässigere und solidere Ergebnisse zu ermöglichen. Diese sollten explizit benannt werden und Anlass für institutionelle Reformen sein, um eine fundiertere Maßnahmengestaltung zu ermöglichen.

Es hat sich grundsätzlich gezeigt, dass die Evaluierungsverfahren von unterschiedlicher Qualität sind, was die Zuverlässigkeit der gewonnenen Erkenntnisse einschränkt. Um eine evidenzbasierte Entscheidungsfindung zu gewährleisten, müssen Evaluierungen eine ausreichende methodische Qualität aufweisen. Die Forschung hat gezeigt, dass in der Regel eine Diskrepanz zwischen der Evaluationstheorie und der praktischen Umsetzung besteht (Huitema et al., 2011). Diese Unterschiede lassen sich auf mehrere Faktoren zurückführen, darunter die Transparenz, Replizierbarkeit und Zuverlässigkeit der angewandten Methoden und der verwendeten Daten.

Interaktionseffekte müssen berücksichtigt werden. Bei der Maßnahmenevaluierung müssen die Wechselwirkungen zwischen einzelnen Energiesparmaßnahmen und den Instrumenten, aus denen die Programme bestehen, berücksichtigt werden (Rosenow et al., 2017). Eine gute Governance umfasst die Evaluierung und Kalibrierung einer großen Anzahl von interagierenden Instrumenten. Zumindest sollte die Berücksichtigung von Wechselwirkungen auf Instrumente ausgedehnt werden, die negative Wechselwirkungen haben und die erwarteten Energieeinsparungen/ THG-Reduzierungen sowohl kurzfristig (statisch) als auch im Zeitverlauf (dynamisch) verringern können. Eine ausgefeiltere Integration von Wechselwirkungen würde auch Folgendes verdeutlichen: (i) positive Wechselwirkungen, bei denen zwei Instrumente Synergieeffekte haben, die größer sind als die Summe ihrer Teile, und (ii) notwendige Wechselwirkungen, bei denen ein ergänzendes Instrument erforderlich ist, um die Bedingungen zu schaffen, die das Hauptinstrument benötigt, um seine Ziele zu erreichen.

Um die Zuverlässigkeit und Vergleichbarkeit von Evaluierungsergebnissen zu gewährleisten, ist eine Standardisierung erforderlich. Ohne Standardisierung der Prozesse verwenden die Bewertenden häufig unterschiedliche Methoden und wenden abweichende Annahmen oder Näherungswerte an (Huitema et al., 2011). Dies erschwert die Verbreitung der Ergebnisse und den Vergleich zwischen verschiedenen Evaluierungen. Eine gute Evaluierungs-Governance sollte darauf abzielen, umfassende und standardisierte Verfahren für die Praxis mit formalen Offenlegungsanforderungen einzuführen (Magro und Wilson, 2017; Schoenefeld und Jordan, 2017).

Bei der Erstellung von Evaluierungen sollte die Erzeugung von Ergebnissen, die zuverlässig, robust, transparent und öffentlich zugänglich sind, im Vordergrund stehen.

Dies ist von wesentlicher Bedeutung für die Verbesserung der Transparenz und der Rechenschaftspflicht bei politischen Gestaltungsoptionen. Um dies zu erreichen, ist es wichtig, dass die Ergebnisse der Evaluierungen in den verschiedenen Veröffentlichungen konsistent sind. Es sollten standardisierte Bewertungsmethoden und Berichtsanforderungen festgelegt werden, um Einheitlichkeit und Vergleichbarkeit zu gewährleisten. Die in den Evaluierungen verwendeten Daten sollten transparent sein und ein klares Verständnis der verwendeten Quellen und Methoden ermöglichen. Außerdem sollten etwaige Lücken

in den verfügbaren Daten anerkannt und klar benannt werden. Es ist von zentraler Bedeutung, dass die Evaluierungsergebnisse öffentlich zugänglich gemacht werden, um eine externe Prüfung und Validierung zu erleichtern. Auf diese Weise kann ein breiterer Kreis von Experten und die Zivilgesellschaft am Evaluierungsprozess teilnehmen, was zu dessen Glaubwürdigkeit und Robustheit beiträgt.

2.3. Der Einsatz von Evaluierungen im politischen Prozess

Wirksame Evaluierungen sowohl der politischen Instrumente als auch staatlicher Kapazitäten für den Vollzug sind für ein wirksames Agenda-Setting sowie eine effektive Durchführung von Reformen zentral. Im Laufe der Zeit kann die Umsetzung von Maßnahmen Rückkopplungseffekte erzeugen, die zu Reformen motivieren können. Dazu gehören Wahrnehmungen darüber, ob eine Maßnahme funktioniert und effektiv ist (kognitive Effekte). Dies kann wiederum motivieren, ob inkrementelle oder radikalere Reformen bestehender Programme (Agenda-Effekte) erforderlich sind (Edmondson et al. 2019). Diese Erkenntnisse hängen in hohem Maße von der Verlässlichkeit und der Evidenz der Evaluierungsprozesse ab, einschließlich der Art und Weise, wie die Evidenz präsentiert und verbreitet wird. Ein weiteres Element des Feedbacks ist die Wahrnehmung, wie gut die mit der Maßnahmengestaltung und -umsetzung beauftragten Verwaltungen arbeiten (Oberlander und Weaver 2015). Verwaltungsfeedback kann den Ausbau staatlicher Kapazitäten (Pierson 1993) oder andere Formen institutioneller Reformen motivieren (Edmondson et al. 2019). Ohne unterstützende Belege ist das Feedback rein diskursiv und normativ, und es fehlen objektive Informationen, auf die sich Evaluierungen stützen können (Schmidt, 2008). In solchen Situationen können unwirksame Elemente des Policy-Mix ohne Prüfung bestehen bleiben. Darüber hinaus werden wirksamere Reformen und Maßnahmenoptionen möglicherweise nicht in Betracht gezogen, da die aktuellen politischen Evaluierungen (fälschlich) zeigen, dass die derzeitigen Programme ausreichend leistungsfähig seien (Weaver 2010). Dementsprechend werden möglicherweise Finanzmittel und Humankapital zur Unterstützung von Programmen eingesetzt, die nicht in dem Umfang und mit der Geschwindigkeit wirken, die zur Erreichung der THG-Einsparungsziele erforderlich sind.

Ex-ante-Evaluierungen und eine fundierte strategische Planung von Maßnahmen setzen die Verfügbarkeit verlässlicher Fakten voraus. Ex-ante-Planungen stützen sich auf Annahmen und die vorhandene Evidenzbasis, die sich aus früher umgesetzten Maßnahmen ergibt. Da das Erreichen von Klimazielen ein langfristiges Unterfangen ist, müssen politische Pfade festgelegt und im Laufe der Zeit Anpassungen vorgenommen werden. Folglich hängt die Genauigkeit dieser Planungen im hohen Maße von der Qualität und Zugänglichkeit von Ex-post-Evaluierungen ab. Ex-post-Evaluierungen spielen eine entscheidende Rolle bei der Beurteilung der Ergebnisse und Auswirkungen vergangener Maßnahmen und liefern somit wertvolle Erkenntnisse für die künftige Planung. Die Verfügbarkeit umfassender und zuverlässiger Ex-post-Evaluierungen ist für die Verbesserung der Effektivität und Präzision strategischer Planungsprozesse zur Erreichung der Klimaziele von entscheidender Bedeutung.

Die Einbeziehung von Evidenz in den Entscheidungsfindungsprozess könnte stärker formalisiert werden. Die Entscheidungsträger könnten verpflichtet werden, sich zur Begründung ihrer Entscheidungen auf verlässlichere Fakten zu stützen, um sicherzustellen, dass die Entscheidungen auf einer soliden Informationsgrundlage beruhen, und um ihre Qualität und Wirksamkeit zu verbessern. Diese formale Einbeziehung von Fakten kann dazu beitragen, Entscheidungsprozesse zu entpolitisieren, indem sie sich auf evidenzbasierte Ansätze stützt, wodurch Entscheidungen weniger von subjektiven Vorurteilen und persönlichen Interessen beeinflusst werden. Darüber hinaus ermöglicht die Einbeziehung von Evidenz ein stärkeres Engagement der Zivilgesellschaft, die dadurch in die Lage versetzt wird, die Entscheidungsträger zu hinterfragen und zur Verantwortung zu ziehen. Dieser partizipatorische Ansatz fördert Transparenz und demokratische Regierungsführung. Die formale Einbeziehung von Evidenz ist von entscheidender Bedeutung, wenn es darum geht, die politischen Entscheidungsträger zur Rechenschaft zu ziehen, da klare Kriterien für die Entscheidungsfindung und das Erfordernis, zuverlässige Evidenz zu verwenden, festgelegt werden können. Auf diese Weise können die Ergebnisse und Auswirkungen politischer Maßnahmen evaluiert und bewertet werden. Insbesondere im Rahmen des Bundes-Klimaschutzgesetzes werden bessere Nachweise sowohl für die Überprüfung und Kontrolle der Fortschritte als auch für den Projektionsbericht benötigt (Ab-

schnitt 4.1.2.). Robuste und umfassende Evidenz ist unerlässlich, um festzustellen, in welchen Bereichen die Politik ihre Ziele möglicherweise nicht erreicht. Auf diese Weise können gezielte Maßnahmen und Kurskorrekturen vorgenommen werden.

Die Motivation für die Durchführung von Evaluationen und deren Beauftragung beeinflusst deren letztendlichen Nutzen. Auch die Nutzung von Evaluationsergebnissen ist pfadabhängig, da Ministerien diese in Auftrag geben, wenn sie nicht intern durchgeführt werden (z.B. an Institutionen wie das Umweltbundesamt (UBA) oder die Deutsche Energieagentur (dena)). Daher kann die Spezifikation dieser Berichte den Umfang, die Qualität und die Verwendung weitgehend mitbestimmen. Wichtig ist die Motivation für die Beauftragung der Berichte. Wenn diese Evaluierungen in erster Linie dazu dienen, die Mittelvergabe eines Ministeriums zur Unterstützung von Programmen zu rechtfertigen, kann es Anreize für methodische Verzerrungen geben, um Kosteneffizienz günstiger darzustellen. Dementsprechend können die Additionalität (Nachweis der Zusätzlichkeit einer Aktivität – ohne die Politikmaßnahme hätte sie nicht stattgefunden) und andere Faktoren, die die Kosteneffizienz beeinträchtigen könnten, aus den Bewertungen ausgeschlossen oder in eher symbolischer Form berücksichtigt werden.

3. Forschungsdesign

Das Forschungsdesign besteht aus drei Schritten. Erstens, eine Literaturrecherche zur Beschreibung des Governance-Rahmens mit einem Schwerpunkt auf Evaluierungen im deutschen Wohnungsbausektor. Zweitens, eine Inhaltsanalyse der Methodikleitfäden und der veröffentlichten, öffentlich zugänglichen Ex-post-Evaluierungsberichte. Schritt 3 schließlich umfasste die Durchführung von Experteninterviews, die dazu beitrugen, die Ergebnisse zu untermauern und die weniger formal kodifizierten Verfahrensaspekte der Durchführung von Evaluierungen und deren Wirkungen zu ermitteln.

Schritt 1: Governance-Rahmen für Wohngebäude in Deutschland und institutionelle Ausgestaltung

Die vorläufige Analyse folgte einem Top-Down-Ansatz, um den Governance-Rahmen im deutschen Wohngebäudesektor zu ermitteln. Dabei lag der Schwerpunkt auf der Rolle der Ex-post-Evaluierung. Dazu wurden zunächst die einschlägigen EU-Richtlinien und die aktuellen Verpflichtungen Deutschlands untersucht. Die Analyse verknüpfte den Governance-Rahmen mit der institutionellen Konfiguration, indem sie die Wechselwirkungen zwischen Verordnungen und Berichtspflichten abbildete. Ziel war es, Gesetze, Strategien und Programme zu kategorisieren, Evaluierungsverfahren und Veränderungen im Laufe der Zeit zu identifizieren sowie Schlüsselthemen und Trends zu beobachten. Darüber hinaus wurde die institutionelle Konfiguration der an der Programmumsetzung und -evaluierung beteiligten Ministerien, Regierungsstellen, Behörden und Beratungsunternehmen kartiert, wobei der Schwerpunkt auf formellen Evaluierungsbeiträgen und öffentlich zugänglichen Berichten lag.

Schritt 2: Inhaltsanalyse von Bewertungsleitfäden und öffentlich zugänglichen Ex-post-Evaluierungen, die nach 2020 veröffentlicht wurden

In einer Inhaltsanalyse wurden Umfang und Qualität bestehender staatlich vorgeschriebener Evaluierungsverfahren bewertet. Die Kodierung fokussierte dabei auf drei Hauptaspekte: den Umfang der verwendeten Evaluierungsmetriken, die zur Bewertung der einbezogenen Metriken angewandten Methoden sowie Datenqualität, Transparenz und Replizierbarkeit. Der Kodierungsansatz umfasste sowohl deduktive als auch induktive Elemente für jede Kategorie.

Hinsichtlich des Geltungsbereichs wurde ein erweitertes Set von Indikatoren aus Tabelle 1 (Abschnitt 2.2) als Bewertungskriterien für „gute Praxis“ verwendet. Die Berichte wurden nach Indikatoren kodiert, um festzustellen, welche von diesen berücksichtigt wurden, und wenn ja, wie sie definiert und angewendet wurden. Für die gemessenen Indikatoren umfasste der Kodierungsprozess die Identifizierung der zur Messung der Variablen verwendeten Metriken und der für ihre Berechnung angewandten Methoden. Diese Messgrößen und Methoden wurden dann im Hinblick auf ihre Transparenz, Reproduzierbarkeit und Qualität kodiert.

Wir kodierten zunächst die Methodikleitfäden und bewerteten dann, wie diese in der Praxis in veröffentlichten Evaluierungen verwendet wurden. Das Verfahren begann mit der Kodierung der Methodikleitfäden und verwendete diese dann als Maßstab dafür, was in öffentlich zugänglichen Evaluierungen enthalten sein sollte. Die veröffentlichten Evaluierungsdokumente wurden dann kodiert, um Unterschiede zwischen den empfohlenen Praktiken und der tatsächlichen Umsetzung zu ermitteln. Dieser Vergleich gibt Aufschluss über die Unterschiede zwischen idealen Praktiken und der Anwendung. Er ermöglicht es uns auch zu beurteilen, inwieweit die Verfahren in den einzelnen Ministerien und Beratungsunternehmen standardisiert sind.

Wir konzentrieren uns auf formelle Berichte, die von Ministerien und beauftragten Beratern als Teil der formellen Evaluierungsverfahren nach der Veröffentlichung der Methodikleitfäden publiziert wurden. Wir beschränken uns auf diese formellen Berichte, die nach 2020 veröffentlicht wurden, da die Leitfäden vor diesem Zeitpunkt weniger transparent und standardisiert waren. Die Veröffentlichung der Leitfäden war eine Reaktion auf dieses Phänomen, so dass wir eine detaillierte Extrapolation der Bewertungen vor ihrer Veröffentlichung als weniger relevant erachteten.

Um die Nachvollziehbarkeit unseres Forschungsdesigns zu verbessern, konzentrieren wir uns auf formelle und öffentlich zugängliche Berichte. Dies gewährleistet ein höheres Maß an Transparenz in unserer Analyse. Die Auswirkungen von nicht in Auftrag gegebenen Maßnahmenevaluierungen und externen Berichten sind weniger nachvollziehbar. Die Ministerien sind nicht verpflichtet, auf diese Veröffentlichungen zu reagieren oder sie anzuerkennen, so dass es spekulativ ist, ihren Nutzen oder ihre Auswirkungen zu bewerten. Darüber hinaus gibt es laufende Evaluierungen, die weniger formalisiert sind und

keiner Berichtspflicht unterliegen, was den Zugang zu Daten für unabhängige Untersuchungen, einschließlich unserer eigenen, erschwert.

Um relevante Dokumente zu ermitteln, haben wir in einem systematischen Ansatz eine umfassende Sammlung relevanter Dokument zusammengestellt. Zunächst durchsuchten wir offizielle Regierungsportale (Ministerien, Behörden, Agenturen), um alle verfügbaren Evaluierungen zu sammeln. Auf diese Weise wurde sichergestellt, dass wir ein breites Spektrum relevanter Dokumente direkt aus maßgeblichen Quellen erfassen konnten. Zweitens haben wir unsere Recherche über die Regierungsportale hinaus auf die Websites von Beratungsunternehmen, Instituten und Nichtregierungsorganisationen (NRO) ausgeweitet. Um den Umfang unserer Suche zu erweitern, nutzten wir außerdem Internet-Suchmaschinen, die es uns ermöglichten, ein breiteres Spektrum an Quellen zu identifizieren (Tabelle 2). Schließlich überprüften wir die zusätzlich gesammelten Dokumente sorgfältig auf mögliche Verweise auf andere Ex-post-Evaluierungen, um eine gründliche Prüfung des verfügbaren Materials zu gewährleisten.

Art der Einrichtung	Anzahl	Beispiele
Bundesregierung	5	Bundesregierung, Bundesministerium für Wirtschaft und Klimaschutz (BMWK), Bundesministerium für Wohnen, Stadtentwicklung und Bauwesen (BMWSB)
Behörde	6	Bundesamt für Wirtschaft und Ausfuhrkontrolle (BAFA), Umweltbundesamt (UBA), Bundesamt für Bauwesen und Raumordnung (BBR)
Beratung	9	Prognos, Guidehouse, PwC
NGO/ Institut	10	Fraunhofer, Öko-Institut e.V., Dena
Andere	4	Expertenrat für Klimafragen, KfW

Tabelle 2: Websites, die für die Inhaltsanalyse nach Evaluierungen durchsucht wurden, n = 34.

Während des Screening-Prozesses haben wir ein breites Spektrum von Dokumenten identifiziert. Die Auswahl der gescreenteten Dokumente (Anhang I) erstreckte sich auf: (a) Strategien, die sich speziell auf die Energieeffizienz beziehen; (b) Ex-ante-Evaluierungen, die Einblicke in die Bewertung politischer Maßnahmen vor ihrer Umsetzung geben; und (c) Nationale Monitoringberichte, die Abschnitte über die Energieeffizienz im Gebäudesektor enthalten und somit relevante Informationen und Analysen bieten.

Aus den gesichteten Dokumenten haben wir unsere Auswahl auf diejenigen eingegrenzt, die sich direkt auf die Erstellung von Ex-post-Evaluierungen beziehen (Tabelle 3). Dazu gehören:

- i. Die Methodikleitfäden für Ex-post-Evaluierungen und Ex-ante-Evaluierungen. Diese Ressourcen halfen uns zu verstehen, wie Ex-post-Evaluierungen durchgeführt werden und in Ex-ante-Evaluierungen nach standardisierten Verfahren einfließen.
- ii. Ex-post-Evaluierungen, die nach der Veröffentlichung der Methodikleitfäden durchgeführt wurden. Diese Evaluierungen boten Einblicke in die Ergebnisse und Auswirkungen der durchgeführten Maßnahmen.
- iii. Abschnitte in den nationalen Überwachungsberichten, die sich speziell mit der Energieeffizienz von Wohngebäuden befassen. Diese Abschnitte lieferten wertvolle Daten und Analysen im Zusammenhang mit unserer Untersuchung.

Bei der Inhaltsanalyse wurde die Kodierung als Hauptverfahren der Datenauswertung eingesetzt. Die Inhaltsanalyse beinhaltet die Kodierung ausgewählter Textstücke nach bestimmten Kodierkategorien (Krippendorff, 2004). Deduktiv begann die Kodierung mit einer Liste von a priori von Edmondson et al. (2023) definierten Herausforderungen für die Gestaltung des Policy-Mix, die sich auf wichtige Bewertungskennzahlen beziehen, die für den Erfolg der Klimapolitik angestrebt werden müssen (Tabelle 1). Wir folgten Mayring (2000) und erstellten zunächst Codebücher in MAXQDA mit einer standardisierten Definition, Indikation, Beispielempfehlung und Kodierregeln für jede Kategorie. Dies verbesserte die Zuverlässigkeit zwischen den Codierern und die Replizierbarkeit unserer Methode.

Zusätzliche Codes wurden induktiv durch die Anwendung des Codebuchs hinzugefügt. Neue Unter-codes wurden induktiv hinzugefügt, während die Methodikleitfäden kodiert wurden, um zu erfassen, wie die Hauptkategorien in der Praxis angewendet wurden (wo dies angemessen war). Dieser schrittweise Prozess kombinierte deduktive und induktive Elemente. Verfahrenstechnisch wurde ein induktives Codebuch erstellt, das Metriken und Methoden in Bezug auf die Anwendung der Bewertungskriterien in den Bewertungsunterlagen enthielt. Die Berichte wurden doppelt kodiert, um die Zuverlässigkeit und Replizier-

barkeit zu verbessern. Während des gesamten Prozesses wurden die Codes iterativ überprüft, zusammengeführt oder aggregiert, um Doppelarbeit zu vermeiden (Krippendorff, 2004).

Typ	Titel	Jahr der Veröffentlichung	Autoren	Auftraggeberin
Ex-post-Evaluierung	Evaluation und Perspektiven des Marktanzreizprogramms zur Förderung von Maßnahmen zur Nutzung erneuerbarer Energien im Wärmemarkt im Förderzeitraum 2019 bis 2020	n.d.	ifeu, Fraunhofer, Fichtner GmbH	BMWK
Ex-post-Evaluierung	Abschlussbericht zur Evaluation der Richtlinie über die Förderung der Heizungsoptimierung durch hocheffiziente Pumpen und hydraulischen Abgleich	2022	Arepo, Wuppertal Institut	BfEE
Formative Ex-post-Evaluierung	Bewertung des Förderprogramms KfW 433	2022	Prognos	BMWK
Formative Ex-post-Evaluierung	Evaluierung der Förderprogramme der EBS WG im Förderzeitraum 2018	2022	Prognos, FIW München	BMWK
Formative Ex-post-Evaluierung	Förderwirkungen BEG EM 2021	2023	Prognos, ifeu, FIW München, iTG	BMWK
Formative Ex-post-Evaluierung	Förderwirkungen BEG WG 2021	2023	Prognos, ifeu, FIW München, iTG	BMWK
Leitfäden	Methodikpapier zur ex-ante Abschätzung der Energie- und THG-Einsparungswirkung von energie- und Klimaschutzpolitischen Maßnahmen	2022	Prognos, Fraunhofer ISI, Öko-Institut e.V.	BMWK
Leitfäden	Methodikleitfaden für Evaluationen von Energie-Effizienzmaßnahmen des BMWi (Projekt Nr. 63/15 - Aufstockung)	2020	Fraunhofer ISI, Prognos, ifeu, Stiftung Umweltenergierecht	BMWK
Nationales Monitoring	Klimaschutz in Zahlen	2022	BMWK	n.a.
Nationales Monitoring	Klimaschutzbericht 2022	2022	BMWK	n.a.
Nationales Monitoring	Zweijahresgutachten 2022 Gutachten zu bisherigen Entwicklungen der Treibhausgasemissionen, Trends der Jahresemissionsmengen und Wirksamkeit von Maßnahmen	2022	Expertenrat für Klimafragen	n.a.
Nationales Monitoring	Die Energie der Zukunft 8. Monitoringbericht zur Energiewende - Berichtsjahre 2018 und 2019	2021	BMWK	n.a.

Nationales Monitoring	Energieeffizienz in Zahlen	2021	BMWK	n.a.
Andere	Begleitung von BMWK-Maßnahmen zur Umsetzung einer Wärmepumpen-Offensive	2023	Dena, Guidehouse, iTG, Öko-Institut, Prognos, EY, pwc, bbh, FIW München, ifeu, heimrich + hannot	BMWK
Strategie	Nationaler Aktionsplan für Energieeffizienz	2014	BMWK	n.a.

Tabelle 3: In die Inhaltsanalyse einbezogene Dokumente. n = 15

Schritt 3: Interviews

In der letzten Phase wurden Expertenbefragungen durchgeführt. Ein Teil unserer Forschungsmotivation liegt in der Untersuchung der Auswirkungen und der Nutzung von Evaluierungen in politischen Entscheidungsprozessen sowie in der breiteren Nutzung von Evidenz in der Entscheidungsfindung. Es sei darauf hingewiesen, dass diese Prozesse oft nicht formalisiert und in der Regel nicht kodifiziert sind, so dass sie einer Inhaltsanalyse nicht zugänglich sind. Daher stützen wir uns auf Interviews als Mittel zur gründlichen Untersuchung dieser Themen. Die Interviews zielen darauf ab, den Einfluss von Evaluierungsprozessen auf die politische Entscheidungsfindung sowie ihre Verbindungen zur Ex-ante-Evaluierung und -Prognose qualitativ zu untersuchen. Obwohl die Entscheidungsfindung ein wesentliches Element der Maßnahmengestaltung ist, räumen wir ein, dass eine systematische Aufschlüsselung dieser typischerweise undurchsichtigen Prozesse den Rahmen unserer Analyse sprengen würde. Darüber hinaus dienen die Interviews dazu, die aus der Inhaltsanalyse und der deskriptiven Darstellung der Daten gewonnenen Erkenntnisse zu ergänzen.

Interview	Akteursgruppe	Rolle/Position	Organisation/Verband	Länge (Min.)
1	Scoping-Interview			45
2	Scoping-Interview			50
3	Politische Entscheidungsträger	Beamte	Bundesministerien	60
4	Politische Entscheidungsträger	Beamte		45
5	Politische Entscheidungsträger	Beamte	Bundesagentur	57
6	Beratung	Evaluator	Konsortium für Evaluation	58
7	Beratung	Evaluator		56
8	Beratung	Evaluator		115
9	Beratung	Evaluator		57
10	Beratung	Evaluator		59
11	Beratung	Evaluator	Unabhängiger Gutachter	58
12	Beratung	Programmleitung	Denkfabrik	65
13	Experte	Akademisch	Monitoringkommission	55
14	Experte	Akademisch	Universität	71

Tabelle 4: Interviewteilnehmende

Wir befragten eine Reihe verschiedener Akteursgruppen, um unterschiedliche Standpunkte zu analysieren. Die für unsere Studie ausgewählte Stichprobe umfasste verschiedene Interessengruppen, darunter Ministerien, Behörden, öffentliche Einrichtungen, Beratungsunternehmen, Experten, Denkfabriken und die Wissenschaft. Zusätzlich zu den Scoping-Interviews wurden 12 Interviews in einem halbstrukturierten Format (Anhang II) mit einer Dauer von jeweils etwa 60 Minuten durchgeführt (Tabelle 4). Unter den Gesprächspartnern befanden sich wichtige Vertretende großer Beratungsunternehmen (Prognos, Fraunhofer ISI, Öko Institut e.V., Guidehouse) sowie Beamte des BMWK und BMWSB. Ein wichtiges Versäumnis war jedoch die Abwesenheit der Bundesförderstellen (KfW und BAFA) in unseren Interviews. Trotz mehrfacher Versuche, ihre Teilnahme sicherzustellen, wurden unsere Anfragen letztlich abgelehnt. Damit ist eine wichtige Akteursgruppe unterrepräsentiert. Dennoch konnten wir den Koordinationsaspekt mit diesen Förderorganisationen durch Gespräche mit anderen Interviewten ansprechen und so die durch ihre Abwesenheit entstandene Lücke teilweise schließen.

Die Interviews wurden transkribiert, gespeichert und mit MAXQDA kodiert. Die Interviews folgten einer ähnlichen Verfahrenslogik in der Kodierungspraxis, wobei auch ein eher induktiver Ansatz zur Erstellung eines Codebuchs verfolgt wurde. Ziel der Interviews

war es, Ansichten zu verfahrenstechnischen Aspekten des Evaluierungsprozesses zu erkunden, einschließlich potenzieller Lücken, Probleme oder Herausforderungen bei der Erstellung von Evaluierungen sowie deren Verbreitung und Nutzung in der Maßnahmengestaltung. Dementsprechend erstreckte sich das Codebuch auch auf institutionelle Vereinbarungen, einschließlich Organisationsstrukturen, Verfahrenslogiken und einschränkender Regeln. Die Codes wurden iterativ überprüft, zusammengeführt oder aggregiert (Krippendorff 2004).

Einschränkungen

Der unterschiedliche Zugang zu den Befragungsteilnehmenden schränkte die Repräsentativität der in unserer Befragungsstichprobe vertretenen Gruppen ein. Unsere Studie stieß aufgrund des eingeschränkten Zugangs zu den Interviewteilnehmenden auf Einschränkungen. Insbesondere war es uns nicht möglich, Vertretende des BAFA und der KfW einzubeziehen, beides wichtige Bundesbehörden, die für die Verwaltung der finanziellen Förderprogramme der Bundesförderung für effiziente Gebäude (BEG) zuständig sind. Obwohl wir Personen aus diesen Organisationen angesprochen haben, wurden unsere Einladungen zur Teilnahme abgelehnt. Während dieses Fehlen eine Lücke in unserer Stichprobe darstellt, haben wir mehrere Teilnehmende aus anderen Akteursgruppen einbezogen, die direkt mit diesen Behörden interagieren. Durch die Triangulation von Informationen aus verschiedenen Quellen und die Ergebnisse unserer Interviews glauben wir, dass unsere Untersuchung auch ohne die Vertretung des BAFA und der KfW gültig ist.

Relativ begrenzte Verfügbarkeit von öffentlich zugänglichen Evaluierungen. Wir untersuchten die Verfügbarkeit von öffentlich zugänglichen Evaluierungen und konzentrierten uns dabei auf Berichte, die nach der Veröffentlichung der Methodikleitfäden im Jahr 2020 publiziert wurden. Die Leitfäden wurden entwickelt, um den Mangel an Standardisierung zwischen den Beratungsunternehmen zu beheben. Daher beinhaltete unsere Analyse keine detaillierte Evaluierung von Berichten vor 2020. Stattdessen konzentrierten wir uns auf: (i) den Inhalt der Leitfäden, um Bereiche zu ermitteln, die noch einer Standardisierung bedürfen, und (ii) die Bestimmung des Ausmaßes, in dem die Leitfäden in der Praxis befolgt wurden. Für die zweite Frage haben wir den Inhalt der veröffentlichten Bewertungen ab 2020 analysiert. Unsere Stichprobengröße war jedoch relativ klein ($n=14$), da in diesem Zeitraum weniger Berichte veröffentlicht wurden.

Verallgemeinerbarkeit einer einzelnen Länderfallstudie. Obwohl diese Analyse auf einer Einzelfallstudie für Deutschland basiert, haben die Ergebnisse Relevanz darüber hinaus und können zum institutionellen Lernen sowie zu Vergleichen mit anderen Ländern beitragen. Die diskutierten Themen stimmen mit der breiteren Literatur über die Evaluierung und Überwachung politischer Maßnahmen überein und ergänzen die bestehende Wissensbasis. Zu diesen Themen gehört etwa die begrenzte Durchsetzung von Energieeffizienzverordnungen, die ein häufig genanntes Problem darstellt und die Glaubwürdigkeit und Wirksamkeit von Energiestandards untergräbt. Unsere Studie unterstreicht auch die Notwendigkeit einer verbesserten vertikalen Koordinierung zwischen den verschiedenen Regierungsebenen und betont die Bedeutung solider methodischer sowie analytischer Ansätze für eine genaue Überwachung der Effekte von Maßnahmen – insbesondere die Bewertung der Auswirkungen auf Treibhausgasemissionen. Diese Themen liefern Erkenntnisse über die Gewinnung und Nutzung von Evidenz in klimapolitischen Prozessen. Weitere Forschungen sind allerdings erforderlich, um die Qualität der Evaluierungen für eine effektive Maßnahmengestaltung und -kalibrierung zu verbessern und die Transparenz sowie Rechenschaftspflicht bei der Entscheidungsfindung zu erhöhen.

4. Institutionelle Ausgestaltung der Governance der Evaluierung von Klimapolitik im Wohngebäudesektor in Deutschland

In Abschnitt 4.1. wird die Ausgestaltung der formalen Verordnung für den Wohnbausektor unter Berücksichtigung der Meldepflichten dargestellt. In Abschnitt 4.2. wird die strukturelle Ausgestaltung der beiden wichtigsten Förderprogramme in diesem Bereich – BEG und GEG (Gebäudeenergiegesetz) – dargestellt. In Abschnitt 4.3. wird dann das Verhältnis zwischen Staat und den mit der Evaluierung in diesem Bereich beauftragten Beratungsunternehmen untersucht.

4.1. Verordnung und Meldepflichten

Die Verordnungs- und Berichterstattungsregelungen für Klimapolitik im Wohngebäudesektor in Deutschland erstrecken sich über mehrere interagierende Verwaltungsebenen. Die EU setzt den regulatorischen Rahmen und hat Berichtspflichten für die Mitgliedsstaaten (Abschnitt 4.1). Deutschland hat ein Klimagesetz, das THG-Ziele für den Gebäudesektor festlegt und durch politische Programme im Auftrag der Bundesministerien umgesetzt wird (Abschnitt 4.2.1.). Im Rahmen des Klima-Aktionsprogramms 2030 werden die Förderprogramme durch Bundesbehörden verwaltet (4.2.1.), während die Vorschriften auf regionaler Ebene erlassen und durchgesetzt werden (Abschnitt 4.2.3.). Weitere Informationen zum Inhalt der jeweiligen Verordnungen sind in Anhang III enthalten.

4.1.1. EU-Richtlinien und Berichterstattung

Die EU gibt den Rahmen für die Mitgliedsstaaten durch Richtlinien vor. Die Richtlinie über die Gesamtenergieeffizienz von Gebäuden (EPBD) sowie die Energieeffizienz-Richtlinien EU/2018/2002 und 2012/27/EU (EED) legen zahlreiche Vorschriften für die Gesamtenergieeffizienz des Gebäudebestands fest. Die Erneuerbare-Energien-Richtlinie EU/2018/2001 (REDII) bietet einen Rahmen für den Ausbau der Erneuerbaren Energien im Wärme- und Kältemarkt auf europäischer Ebene. Unter anderem verpflichtet Artikel 14 Absatz 1 EED die Mitgliedsstaaten, alle fünf Jahre einen Bericht über die Wärme- und Kühleffizienz zu erstellen.

Artikel 7 der EPBD (bestehende Gebäude) verlangt von den Mitgliedstaaten eine Erklärung, wo sie von den EPB-Standards abweichen. Die neue EPBD zwingt die Mitgliedstaaten nicht, die EPB-Standards anzuwenden. Sie verlangt aber von Mitgliedstaaten zu erklären, wo und warum sie ggf. von diesen Standards abweichen. Diese Berichterstattung soll die Umsetzung in den Mitgliedstaaten vorantreiben.

Die Berichterstattungsanforderungen der nationalen Energie- und Klimapläne (NECP) beruhen in erster Linie auf nationalen Ex-ante-Evaluierungen. Die nationalen Energie- und Klimapläne (NECP) wurden durch *die Verordnung über die Governance der Energieunion und Klimamaßnahmen* (EU) 2018/1999 eingeführt. Die letzte Berichterstattung wurde im Juni 2020 in den NECP aufgenommen.

Die Berichtspflichten werden durch die EU-Governance-Richtlinie weiter konsolidiert. Künftig wird die Europäische Union (EU) die Berichtspflichten auf europäischer Ebene durch die EU-Governance-Verordnung (Richtlinie 2018/1999) konsolidieren, die Ende 2018 verabschiedet wurde. Diese Konsolidierung umfasst verschiedene Verpflichtungen, wie die Nationalen Energie- und Klimapläne (NECP), die EU-Energieeffizienzrichtlinie (EED; Richtlinie 2012/27/EU und überarbeitet 2018/844), die EU-Leitfäden für die Bewertung staatlicher Beihilfen und die europäischen Standards für die CO₂-Überwachung und Berichterstattung.

Die EU-Berichterstattung beeinflusst die Umsetzung der nationalen Politik, sieht aber keine Sanktionen für die Nichteinhaltung vor. Die EU-Berichterstattungsanforderungen haben Einfluss auf die Maßnahmengestaltung und Entscheidungsfindung und haben die Berichterstattungsverfahren in hohem Maße geprägt. Sie sehen jedoch keine Sanktionen für die Nichteinhaltung vor und führen daher zu einer geringeren Rechenschaftspflicht der nationalen Regierung, als wenn eindeutige Sanktionen festgelegt würden oder die Verpflichtungen rechtsverbindlich wären.

4.1.2. Sektorübergreifende Verordnung

Bundes-Klimaschutzgesetz (KSG)

Das KSG enthält ein rechtsverbindliches Sektorziel: Die Treibhausgasemissionen im Gebäudesektor müssen bis 2030 um 68 % gegenüber 1990 reduziert werden, wobei die jährlichen Ziele in den 2020er Jahren linear abnehmen. Das revidierte Bundes-Klimaschutzgesetz (KSG) legt die Verpflichtung der Bundesregierung fest, die Treibhausgasemissionen bis 2030 um 65 % gegenüber 1990 zu senken und bis 2045 Treibhausgasneutralität zu erreichen. Das KSG legt auch spezifische Ziele für verschiedene Sektoren fest. Im Gebäudesektor sollen bis 2030 67 Mio. t CO₂eq. erreicht werden, was einer Reduktion von 68 % gegenüber 1990 entspricht.

Die Klimaziele werden durch ein kontinuierliches Monitoring überprüft. Alle zwei Jahre legt der Expertenrat für Klimafragen (ERK) einen Bericht über die erreichten Ziele sowie Maßnahmen und Trends vor. Der erste Zweijahresbericht wurde im Jahr 2022 erstellt (ERK 2022a). Der Bericht unterliegt der Berichtspflicht gegenüber dem Bundestag, der EU und den Vereinten Nationen. Nationale Berichtspflichten haben mehr Gewicht als die auf EU-Ebene, die nicht verbindlich sind. Wenn sich Instrumente nicht als wirksam erweisen, müssen sie neu formuliert oder ersetzt werden. Der Rat bereitet auch Bewertungen zu anderen Aspekten des Bundes-Klimaschutzes vor. Der Rat hat die Auswirkungen des Sofortprogramms auf die Sektoren Gebäude und Verkehr geprüft (ERK 2022b). Zuletzt wurde am 17.04.2023 das Gutachten zu den Treibhausgasemissionen Deutschlands für das Jahr 2022 veröffentlicht (ERK 2023).

Die Bundesregierung sieht den Projektionsbericht als zentralen Kontrollmechanismus für ihre Klimaschutzpolitik. Die Folgenabschätzung ist durch das KSG [§ 9 (2) KSG] vorgeschrieben und wird dem Bundestag und dem Bundesrechnungshof (BRH) berichtet (BRH 2022: 38). Im November 2021 hat die Bundesregierung den Projektionsbericht 2021 für das Jahr 2020 verabschiedet. Dieser Bericht enthält erstmals eine Prognose (Ex-ante-Evaluierung) der zu erwartenden Minderungswirkung der aktuellen Klimaschutzmaßnahmen (BRH 2022: 39). Der nächste Bericht (Projektionsbericht 2023) war zum Zeitpunkt des Abschlusses dieses Berichts (Juli 2023) noch in Arbeit.

Verordnung	Kategorie (Instrumente)	Sektor(en)	Monitoring	Frequenz/Jahr	Berichterstattung	Format (Methodik)	
EU	Nationale Energie- und Klimapläne (NECP)	Richtlinie	Sektorenübergreifend	NECP-Berichterstattung	Alle 3 Jahre	EU	Ex-ante
	Richtlinien zur Energieeffizienz (EED)	Richtlinie	Gebäude	Artikel 14	Alle 5 Jahre	EU	Ex-ante
	Richtlinie über die Gesamtenergieeffizienz von Gebäuden (EPBD)	Richtlinie	Gebäude	Artikel 7	Alle 2 Jahre	EU	Ex-ante
National	Bundes-Klimaschutzgesetz (KSG)	Klimagesetz	Sektorenübergreifend	Abschnitt 12 (1)	Jährlich	Bundestag	Hybrid (Aggregation)
				Bericht über die Projektionen	Alle 2 Jahre	UN, EU, Bundestag	
	Klima-Aktionsprogramm 2030	Strategie	Sektorenübergreifend	Bericht über die Überwachung	Jährlich	Bundestag, BRH	Hybride (Aggregation)
				Fortschrittsbericht	Alle 3 Jahre		
	Energieeffizienz-Strategie 2050	Strategie	Sektorenübergreifend	Zwischenbericht	2020	Bundestag, BRH	„Hybrid“ (Aggregation)
				Bericht über die Überwachung	2022		
	Nationaler Aktionsplan für Energieeffizienz (NAPE)	Programm (Aggregat)	Sektorenübergreifend	NAPE-Überwachung	Jährlich	Bundestag, BRH	Ex-post (Aggregation)
	Das Treibstoff-Emissionshandelsgesetz (BEHG)	Programm (CO ₂ Steuer)	Gebäude/Transport	Fortschrittsbericht	Alle 2 Jahre	Bundestag	Ex-post (top-down)
Bundesförderung für energieeffiziente Gebäude (BEG)	Programm (Finanzierung)	Gebäude	Berichterstattung	Jährlich	Bundestag, BRH	Formative Ex-post (bottom-up)	
			BEG-Auswertung	Alle 2 Jahre		Ex-post (bottom-up)	
			BEG-Auswertung	Alle 4 Jahre	Bundestag, BRH	Ex-post (bottom-up)	
Gebäudeenergiegesetz (GEG)	Programm (regulatorisch)	Gebäude	K.A.	K.A.	K.A.	K.A.	

Tabelle 5: Formelle Institutionen (Verordnungen) im Wohngebäudebereich, zusammen mit den Anforderungen an die Überwachung, Berichterstattung und Bewertung. Die rot hervorgehobene Zelle zeigt an, dass das BEG die einzige Verordnung ist, das durch eine Bottom-up-Ex-post-Methode evaluiert wird. Die orangefarbene Einfärbung zeigt an, dass keine veröffentlichten Bewertungen verfügbar sind.

Die Prüfung durch den Bundesrechnungshof (BRH) zeigt, dass der Klimaschutzbericht seine Monitoringziele verfehlt. Dem Bericht fehlen nach Ansicht des BRH wichtige Informationen, wie zum Beispiel die THG-Einsparungen, die die Bundesregierung von den einzelnen Klimaschutzmaßnahmen erwartet bzw. bisher mit ihnen erreicht hat (BRH 2022: 37). Nach Angaben des BRH enthielten die bisherigen Klimaschutzberichte keine Angaben zu den erzielten Wirkungen der aktuellen Klimaschutzmaßnahmen, die entsprechenden Daten lägen der Bundesregierung aber vor.

Aktuelle Projektionen zeigen, dass die bestehenden Klimaschutzmaßnahmen nicht ausreichen, um diese ehrgeizigen gesetzlichen Ziele zu erreichen. Folglich wird für Deutschland im Jahr 2030 eine Minderungslücke von 195 Mio. t CO₂ eq prognostiziert, was 27 % der Gesamtemissionen im Jahr 2020 entspricht. Im Gebäudesektor werden die Emissionen schätzungsweise nur um etwa 57 % sinken, was einer Lücke von 24 Mio. t CO₂ eq entspricht (Umweltbundesamt 2022).

Nationale Strategien/Programme

Das Klimaschutzprogramm spezifiziert den nationalen, sektorübergreifenden klimapolitischen Instrumentenmix. Erstmals 2014 als Aktionsprogramm Klimaschutz 2020 eingeführt, wurde 2019 das Klimaschutzprogramm 2030 verabschiedet. Das Klimaschutzprogramm 2030 umfasst mehrere Arten von politischen Instrumenten, darunter Förderprogramme, Steuererleichterungen und Regulierungsmaßnahmen.

Der jährliche Monitoring-Bericht ist das Kernstück des Monitoring-Prozesses für die Energiewende. Alle drei Jahre wird anstelle des Monitoringberichts der ausführlichere Fortschrittsbericht zur Energiewende vorgelegt. Am 3. Dezember 2014 hat die Bundesregierung zum ersten Mal einen solchen Fortschrittsbericht veröffentlicht. Mit diesem Bericht erfüllt die Bundesregierung gleichzeitig ihre Berichtspflichten nach § 63 Abs. 1 Energiewirtschaftsgesetz (EnWG), § 98 Erneuerbare-Energien-Gesetz (EEG) und § 24 der Energiemarktstammdatenregisterverordnung (MaStRV) sowie nach dem Nationalen Aktionsplan Energieeffizienz (NAPE) und der Energieeffizienzstrategie für Gebäude (ESG). Dieses Berichtsverfahren verdichtet eine große Menge an verfügbaren energiestatistischen Informationen. Bereits umgesetzte Maßnahmen fließen ebenso in die Analyse ein wie die Frage, in welchen Bereichen in Zukunft Anstrengungen erforderlich sind.

Eine Kommission unabhängiger Energieexperten überwacht den Monitoringprozess.

Auf der Grundlage wissenschaftlicher Erkenntnisse nehmen die Mitglieder der Kommission anschließend zu den Monitoring- und Fortschrittsberichten der Bundesregierung Stellung. Der Bericht wird im Rahmen von zwei Forschungsprojekten erstellt, die von der Bundesstelle für Energieeffizienz (BfEE) beim Bundesamt für Wirtschaft und Ausfuhrkontrolle (BAFA) und dem Umweltbundesamt (UBA) betreut werden.

Nationale Programme sind prüfungspflichtig. Nationale Ausgabenprogramme unterliegen der Berichterstattung und Evaluierung durch den Bundesrechnungshof (BRH), die oberste Bundesbehörde für Finanzkontrollen der Bundesrepublik Deutschland. Bei allen Instrumenten, die für die öffentliche Finanzierung auf nationaler Ebene relevant sind, überwacht der BRH die Erfolgskontrolle dieser Programme, einschließlich der Frage, wie die Mittel ausgegeben werden und wie effizient die Förderprogramme sind. Er fordert jährlich Zahlen zu allen finanzierungsrelevanten Programmen, die durch die Beauftragung von Ex-post-Evaluierungen erhoben werden [Interview 7, 8]. Dies wird seit dem Aktionsprogramm Klimaschutz 2020 durchgeführt, das auch den Nationalen Aktionsplan Energieeffizienz im Jahr 2014 beinhaltetete, als diese Prozesse eingeführt wurden [Interview 7, 8].

Probleme bei der Überwachung wurden vom BRH in früheren Bewertungszeiträumen festgestellt. Bei der Bewertung des Aktionsprogramms Klimaschutz 2020 durch den BRH wurden erhebliche Probleme beim Monitoring festgestellt, da die THG-Einsparungseffekte der einzelnen Maßnahmen nicht spezifiziert sind und ein großer Teil der Maßnahmen nicht direkt zu einer Reduktion beiträgt. Das BMUV begründete seine Position damit, dass der Projektionsbericht (alle 2 Jahre) aktuelle Entwicklungen, wie das Klimaschutz-Sofortprogramm 2022 oder die Verteuerung der Zertifikate im EU-ETS, seit Anfang 2021 nicht berücksichtige.

Klimaschutzprogramm 2030

Im Klimaschutzprogramm 2030 ist das Ziel verankert, die Emissionen aus dem Gebäudesektor auf 72 Millionen Tonnen CO₂ pro Jahr zu reduzieren (BReg2019). Der Instrumentenmix besteht aus erhöhten Subventionen, CO₂ Preisen und regulatorischen Maßnahmen (BReg2019). Außerdem wurden Steuervergünstigungen für energieeffiziente Ge-

bäudesanierungen eingeführt. Energieeffiziente Sanierungsmaßnahmen, wie der Austausch von Heizungsanlagen, der Einbau neuer Fenster sowie die Dämmung von Dächern und Außenwänden, sollen ab 2020 steuerlich gefördert werden. Gebäudeeigentümer aller Einkommensklassen werden durch einen Steuerabzug gleichermaßen profitieren. Die Fördersätze der bestehenden KfW-Förderprogramme wurden um 10 % erhöht (BReg2019).

Das Klimaschutzprogramm 2030 enthält 96 sektorale und sektorübergreifende Maßnahmen zur Emissionsminderung. Das Programm enthält keine Zielwerte für die THG-Einsparung der einzelnen Maßnahmen (BRH 2022: 17). Das Programm führt einige Maßnahmen des Programms 2020 fort, die nachweislich nicht zu einer THG-Einsparung beigetragen haben (BRH 2022: 17).

Energieeffizienz-Strategie 2050

Die deutsche Energieeffizienzstrategie 2050 dient als Rahmen für die Verbesserung der Energieeffizienzpolitik in Deutschland. Damit orientiert sie sich am Energieeffizienzziel der Europäischen Union, den Primär- und Endenergieverbrauch bis 2030 um mindestens 32,5 % zu senken. Die Strategie legt ein Energieeffizienzziel für 2030 fest und fasst die erforderlichen Maßnahmen in einem neuen Nationalen Aktionsplan für Energieeffizienz (NAPE 2.0) zusammen. Darüber hinaus enthält sie Leitfäden dafür, wie der Dialogprozess für den Fahrplan für Energieeffizienz 2050 strukturiert werden sollte, um eine effektive Einbindung und Zusammenarbeit der Interessengruppen zu fördern.

Die Energieeffizienzstrategie wird an die EU in Erfüllung der deutschen NECP-Anforderungen und Art. 7 EED kommuniziert. Im Jahr 2023 wird die EU prüfen, ob die europäischen Reduktionsziele erhöht werden müssen. Auch in Deutschland ist die Erstellung eines Monitoring-Berichts geplant. Darin soll untersucht werden, ob das Effizienzziel für 2030 im Hinblick auf das langfristige Ziel der Treibhausgasneutralität noch angemessen ist, oder ob es verschärft werden muss. Die Bundesstelle für Energieeffizienz (BfEE) unterstützt das BMWK bei der Umsetzung der Roadmap Energie-Effizienz 2050. Die Umsetzung des Roadmap-Prozesses wird von einem Konsortium um Prognos begleitet. Eine vom BMWK eingesetzte wissenschaftliche Begleitgruppe stellt die Einbindung der Forschungsplattformen des BMWK in den Roadmap-Prozess sicher. Die administrativen Aufgaben werden von der Deutschen Energie-Agentur (dena) wahrgenommen.

Nationaler Aktionsplan für Energieeffizienz

Der NAPE (Nationaler Aktionsplan Energieeffizienz) wurde als umfassendes Maßnahmenpaket zur Verbesserung der Energieeffizienz in Deutschland konzipiert. Der NAPE wurde erstmals 2014 zusammen mit dem Klimaaktionsplan 2020 umgesetzt. Der NAPE 2.0 wurde im Jahr 2019 als Teil der Energieeffizienzstrategie 2050 (BMWi 2019) umgesetzt. Er berücksichtigt politische Erkenntnisse und passt sich an neue Entwicklungen an, insbesondere mit Blick auf den Zeitraum von 2021 bis 2030. Der NAPE 2.0 wird in diesem Jahr in Anlehnung an den neuen Fahrplan „Energieeffizienz für eine klimaneutrale Zukunft 2045“ erneut aktualisiert.

Der NAPE muss jährlich überwacht werden. Der NAPE wird seit 2014 im Auftrag der BfEE jährlich überwacht. Der letzte öffentlich zugängliche Bericht wurde jedoch im Jahr 2021 für den Berichtszeitraum 2018-2019 veröffentlicht (NAPE-Monitoring-2021). Es ist unklar, ob dieses Ausbleiben der Veröffentlichung auf den Regierungswechsel, COVID, den Ukraine-Konflikt oder eine Kombination all dieser Faktoren zurückzuführen ist.

Die aggregierte Bewertung fließt in die NAPE-Berichterstattung ein. Die NAPE-Berichterstattung ist der wichtigste Mechanismus zur Bewertung von Energieeffizienzmaßnahmen. Es handelt sich um einen umfangreichen sektorübergreifenden Bericht, der auch den Gebäudesektor umfasst. Die aggregierte NAPE-Berichterstattung stützt sich auf Bottom-up-Ex-post-Evaluierungen, Modellierung und statistische Extrapolation.

4.1.3. Sektorale Programme

Das Brennstoffemissionshandelsgesetz (BEHG)

Mit dem Brennstoffemissionshandelsgesetz (BEHG) wurde ein eine CO₂-Bepreisung für Gebäude und Verkehr bis 2026 eingeführt. Mit dem BHEG wurden alle Brennstoffemissionen außerhalb des EU-Emissionshandelssystems (EU ETS) ab 2021 in das nationale Emissionshandelssystem (nEHS) einbezogen. Von 2021 bis 2025 funktioniert das nEHS als Emissionshandelssystem mit festen CO₂-Preisen, die jährlich steigen, beginnend mit

25 Euro/tCO₂ im Jahr 2021 und bis 45 Euro/tCO₂ im Jahr 2025. Ab 2026 wird ein Preiskorridor von 55 bis 65 Euro/tCO₂ eingeführt, wobei die Zukunft des BHEG über 2026 hinaus noch unklar ist.

Das BEHG sieht Berichtspflichten und Revisionschritte alle zwei Jahre vor, beginnend im Jahr 2022. Die Bundesregierung ist verpflichtet, Evaluierungen des Gesetzes durchzuführen (BMUV 2021). Diese Evaluierungen sind als Fortschrittsberichte bis zum 30. November 2022 und 30. November 2024 dem Bundestag vorzulegen. Danach sollen die Evaluierungen alle vier Jahre durchgeführt werden. Die Fortschrittsberichte müssen sich auf den Umsetzungsstand und die Wirksamkeit des nationalen Emissionshandelssystems konzentrieren. Sie sollen sich auch mit den Auswirkungen der in § 10 Abs. 2 des Gesetzes genannten Festpreise und Preiskorridore befassen. Auf der Grundlage dieser Erkenntnisse schlägt die Regierung notwendige Gesetzesänderungen vor, um das Emissionshandelssystem anzupassen und weiterzuentwickeln. Darüber hinaus ist die Regierung verpflichtet, die jährlichen Klimaschutzberichte nach § 10 des Bundes-Klimaschutzgesetzes in diesem Prozess zu berücksichtigen.

Im ersten Revisionsschritt wurde der Preispfad aufgrund der Volatilität der Gaspreise und des Drucks auf die Heizkosten durch Handelsschocks infolge des Ukraine-Konflikts pausiert. Im Rahmen des dritten Entlastungspakets der Bundesregierung Anfang 2022 beschloss der Koalitionsausschuss, die geplanten Preiserhöhungen im nationalen Emissionshandel (nEHS) ab 2023 um jeweils ein Jahr zu verschieben. Diese Entscheidung wurde durch eine Novelle des BEHG umgesetzt, die im November 2022 in Kraft trat.

Bundesförderung für energieeffiziente Gebäude (BEG)

Ein wichtiger Baustein des Aktionsprogramms Klimaschutz 2030 ist die Bundesförderung für energieeffiziente Gebäude (BEG). Die BEG bündelt die bestehenden Gebäudeförderprogramme (CO₂-Gebäudemodernisierungsprogramm, Marktanzreizprogramm (MAP), Anreizprogramm Energieeffizienz (APEE) und Förderprogramm Heizungsoptimierung (HZO)) ab 2021 in einem neuen System, das auf die Bedürfnisse der Zielgruppen ausgerichtet ist (BMWi 2021: 92).

Das BEG gliedert sich in drei Teilprogramme: das BEG Wohngebäude (BEG WG), das BEG Nichtwohngebäude (BEG NWG) und das BEG Einzelmaßnahmen (BEG EM). Weitere

Details zur Unterscheidung zwischen diesen Teilprogrammen finden Sie im Anhang III. Träger der Förderprogramme bleiben die KfW und das BAFA (Bundesamt für Wirtschaft und Ausfuhrkontrolle) (Abschnitt 4.2.1.).

Die erste formale Evaluierung des BEG wurde kürzlich veröffentlicht (Juni 2023). Die Evaluierung wurde von Prognos koordiniert, die Berichte zu den drei aggregierten Teilprogrammen wurden unabhängig voneinander veröffentlicht. Aggregierte Daten und Leistungskennzahlen zur BEG wurden zuletzt im dritten Quartal 2021 auf der BMWK-Website veröffentlicht (BMWK 2021).

Gebäudeenergiegesetz (GEG)

Das Gebäudeenergiegesetz (GEG) stellt einen wichtigen Baustein des Klimaschutzprogramms 2030 dar. Das Gebäudeenergiegesetz (GEG) ist am 1. November 2020 in Kraft getreten (BMWi 2021: 92) und wurde zum 01.01.2023 novelliert. Ob und inwieweit die Ziele des Energiekonzepts mittel- bis langfristig erreicht werden können, wird die Bewertung zeigen. Sie soll auch aufzeigen, welche neuen Maßnahmen ergriffen werden müssen.

Das GEG stellt Anforderungen an bestehende Gebäude für Nachrüstungen und Sanierungen. Das GEG stellt für bestimmte Teile (Austausch bestimmter alter Heizkessel, Dämmung bestimmter Rohrleitungen, Dämmung der obersten Geschossdecke, Einbau bestimmter Regelungstechnik von Heizungs- und Klimaanlage) maßnahmeunabhängige Nachrüstungsanforderungen. Es legt auch Standards für Renovierungsarbeiten fest, die ansonsten zur Einhaltung der Energiestandards durchgeführt werden müssten.

Die nach Landesrecht zuständige Behörde kann auf Antrag auch Ausnahmen von den bedingten Anforderungen zulassen. Diese gelten insbesondere bei mangelnder Wirtschaftlichkeit im Einzelfall. Allerdings sind die Vorgaben, wie diese Ausnahmen anzuwenden sind, nur lose definiert.

4.2 Governance und Ministerien

In diesem Abschnitt wird die institutionelle Konfiguration von Ministerien, Bundesbehörden und Beratungsunternehmen skizziert, die zu den derzeitigen Verfahren für Evaluierungen im deutschen Wohngebäudesektor beitragen (Abbildung 1).

Bei der Erörterung dieser Organisationen werden die ministeriellen Mandate, die Koordination und die Rolle in den Evaluierungsverfahren umrissen. Obwohl es eine Koordination über die gesamte institutionelle Struktur hinweg gibt, ordnen wir für diesen Bericht die Institutionen entsprechend der Durchführung der beiden wichtigsten Programme an: (i) Bundesförderung für effiziente Gebäude (BEG) und (ii) das Gebäudeenergiegesetz (GEG). Dieser strukturelle Aufbau wird weiter in nationale Ministerien und Behörden sowie in subnationale dezentralisierte lokale Behörden eingeteilt. Andere relevante Akteure werden auch im Hinblick auf ihre Rolle bei der Bewertung dieser Programme skizziert.

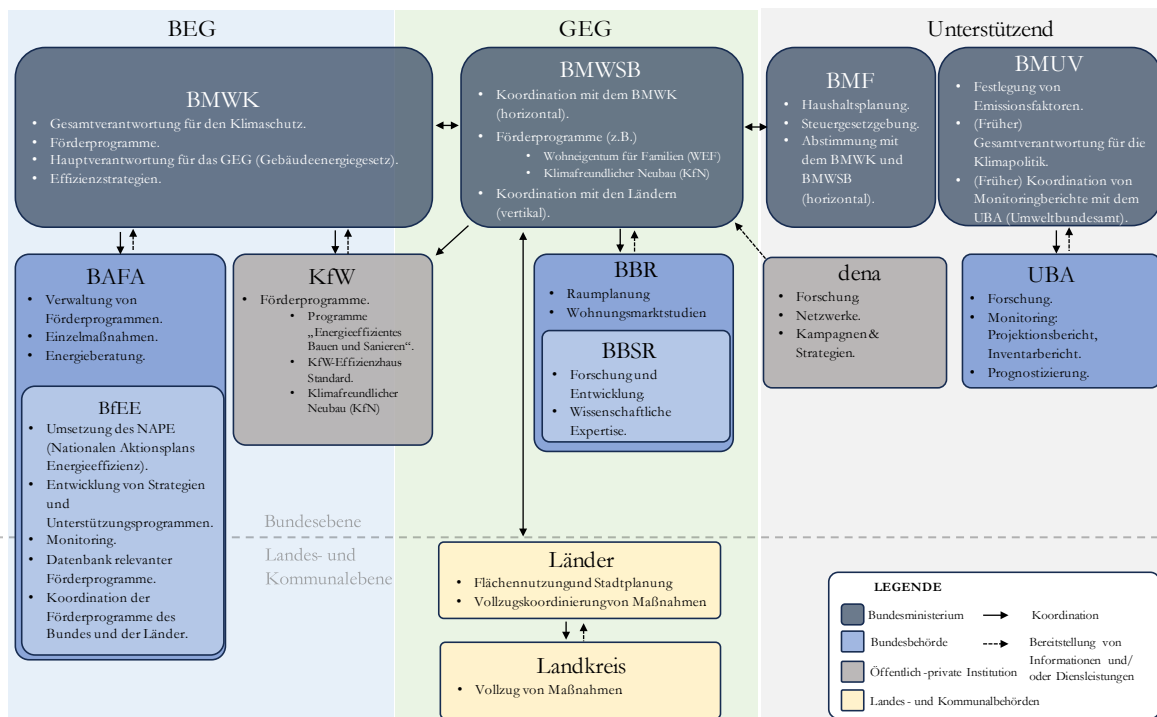


Abbildung 1: Strukturelle bürokratische Einrichtungen für die Durchführung und Bewertung des BEG und des GEG. Quelle: Eigene Darstellung

4.2.1. Bundesfördermittel für effiziente Gebäude (BEG)

Bundesministerien

Das Bundesministerium für Wirtschaft und Klimaschutz (BMWK) hat die Gesamtverantwortung für die Energie- und Klimapolitik und beaufsichtigt die Umsetzung des BEG.

Das BMWK hat ein breites Mandat, das verschiedene Aufgaben umfasst. Dazu gehören die Umsetzung der Energieeffizienzstrategie, die Überwachung mehrerer Förderprogramme einschließlich des BEG, die Verwaltung der Förderinitiative *EnEff.Gebäude.2050* für wichtige Projekte und das Monitoring der Fortschritte der Energiewende. In dieser Funktion arbeitet das BMWK mit den Fördergebern des Bundes zusammen und ist hauptverantwortlich für die Initiierung von Evaluierungen und Bewertungen durch ein Netzwerk von Beratungsunternehmen (Abschnitt 4.3.). Darüber hinaus betreut das BMWK die Monitoring-Kommissionen zur Energiewende. Das Gesamtprogramm der BEG wurde kürzlich (Juni 2023) für den Zeitraum Q3 2021 evaluiert und auf der Website des BMWK veröffentlicht. Ab 2022 läuft derzeit eine umfragebasierte Evaluierung des BEG¹, die vom Beratungsunternehmen Prognos² koordiniert wird.

Die BEG-Evaluierung ist als begleitende/formative Evaluierung konzipiert. Dies liegt vor allem daran, dass die vollständige Wirkung erst mit einer Verzögerung von drei bis vier Jahren (Bereitstellung/Aufruffrist bis Vorlage/Prüfung des Verwendungsnachweises) festgestellt werden kann. Dies bedeutet, dass zum Zeitpunkt der Evaluierung ein Teil der bewilligten/geförderten Projekte noch nicht umgesetzt ist, sondern nur die Antragsdaten vorliegen. Methodisch wird diesem Aspekt durch die Einbeziehung der Storno-/Verzichtsquote Rechnung getragen. Ähnlich verhält es sich bei der KfW-433-Evaluierung. In beiden Evaluierungen ist bzw. war das jeweilige Programm noch nicht abgeschlossen [Interview 7]. Eine Besonderheit der BEG war die Unterstützung des BMWK aus der laufenden Evaluierungsarbeit bei der Entwicklung der Richtliniennovelle Sommer 2022 und für die Richtlinie 2023 sowie bei anderen politischen Prozessen im betreffenden Themenfeld [Interview 7]. Bei der Evaluation der EBS AG waren die Informationen über den Förderoutput hingegen „verlässlicher“, da die Verzögerung deutlich geringer war. Diese Evaluierung

¹ <https://www.energiewechsel.de/KAENEF/Redaktion/DE/Meldungen/2022/2022-03-31-befragung-fuer-evaluation-der-beg-startet.html>

² <https://www.prognos.com/de/FAQ-BEG>

entspricht daher eher einer Ex-post-Evaluierung; auch weil die EBS WG zum Zeitpunkt der Evaluierung abgeschlossen und an die BEG übergeben war [Interview 7].

Bundesbehörden

Die BEG-Förderung wird durch Bundesstellen koordiniert. Das BMWK hat sich zusammen mit den Ministerien für Wohnen, Stadtentwicklung und Bauwesen (BMWSB) und Finanzen (BMF) auf ein gemeinsames Vorgehen bei der Bundesförderung für effiziente Gebäude (BEG) verständigt. Diese wird mit den Bundesbehörden, dem Bundesamt für Wirtschaft und Ausfuhrkontrolle (BAFA) und der Kreditanstalt für Wiederaufbau (KfW), koordiniert und von diesen durchgeführt. Diese Ministerien und Agenturen arbeiten zusammen, um Finanzierungsinitiativen und Programme zur Förderung der Energieeffizienz im Bausektor bereitzustellen. Die Evaluierung der einzelnen Zweige des BEG wird durch die jeweiligen Bundesagenturen und beauftragte Berater koordiniert.

Das Bundesamt für Wirtschaft und Ausfuhrkontrolle (BAFA) ist zuständig für die Verwaltung von Fördermitteln für die Sanierung von Wohngebäuden, für Neubauten und für Energieberatungsleistungen. Die *Abteilung 5 (Energieeffizienz, Erneuerbare Energien, Besondere Ausgleichsregelung)* ist zuständig für Aufgaben und Förderprogramme im Bereich der Erneuerbaren Energien und der Energieeffizienz. Die *Abteilung 6 (Klimaschutzgebäude, Energie-Info-Center, Anpassungsgeld)* ist für Förderprogramme im Bereich der Erneuerbaren Energien zuständig. Die Energieberatung umfasst die Abwicklung der Förderverfahren (Antragstellung, Verwendungsnachweis), die Bundesförderung für Energieberatung für Wohngebäude (EBW) einschließlich iSFP-Programm, durchgeführt von PricewaterhouseCoopers GmbH, und die Zulassung von Energieberatern.

Das Bundesstelle für Energieeffizienz (BfEE) unterstützt das BMWK konzeptionell in allen Fragen der Energieeffizienz. Die BfEE ist eine Unterabteilung des BAFA, die für die Umsetzung von Maßnahmen des Nationalen Aktionsplans Energieeffizienz (NAPE), die Entwicklung der Effizienzstrategie Gebäude und die Entwicklung neuer Förderprogramme zuständig ist. Das Zentrum führt auch PR-Kampagnen zur Förderung der Energieeffizienz durch. Im Rahmen des Monitorings ist sie verantwortlich für die Ermittlung der Energieeinsparungen in Deutschland und die entsprechende Berichterstattung sowie für die Beobachtung und Bewertung der Energiedienstleistungsmärkte mit dem Ziel, diese zu entwickeln. Der Monitoring-Prozess zur Energiewende „Energie der Zukunft“

dient dazu, die Umsetzung des Energiekonzepts und des Maßnahmenprogramms der Bundesregierung zu überprüfen und bei Zielverfehlungen gegenzusteuern. Der Bereich des Energieverbrauchs und der Energieeffizienz ist ein Schwerpunkt des Monitoring-Prozesses. Darüber hinaus wird über den Ausbau der Erneuerbaren Energien, die Treibhausgasemissionen, die Versorgungssicherheit, die Netzinfrastruktur und die Energiewende im internationalen Kontext berichtet.

Öffentlich-private Einrichtungen

Die Kreditanstalt für Wiederaufbau (KfW) ist mitverantwortlich für die Verwaltung des BEG-Förderprogramms. Seit der Einführung der Bundesförderung für energieeffiziente Wohngebäude (BEG WG) und Neubauten (BEG NWG) am 1. Juli 2021 wird der Gebäudestandard als „KfW-Effizienzhaus“ bezeichnet. Bei der Sanierung fördert die KfW verschiedene Effizienzhausstandards, darunter Denkmal, 100, 85, 70, 55 und 40 (Tabelle 6). Bei Neubauten wurden die Effizienzhausstufen 55 und 40 gefördert.

KfW-Effizienzstandard	Energieverbrauch
KfW-Effizienzhaus 100	Verbraucht genauso viel wie ein neues GEG-Referenzgebäude.
KfW-Effizienzhaus 85	Benötigt nur 85 Prozent der Energie eines neuen GEG-Gebäudes.
KfW-Effizienzhaus 70	Benötigt nur 70 Prozent des Energiebedarfs eines neuen GEG-Gebäudes.
KfW-Effizienzhaus 55	Benötigt nur 55 Prozent der Energie eines neuen GEG-Gebäudes.
KfW-Effizienzhaus 40	Benötigt nur 40 Prozent der Energie eines neuen GEG-Gebäudes.

Tabelle 6: KfW-Effizienzstandards für Sanierungen.

Die Förderung des KfW 55-Standards für Neubauten wurde kürzlich gestrichen und durch das Programm Klimafreundlicher Neubau (KfN) ersetzt. Die Förderung des Effizienzhaus 55-Standards für Neubauten ist komplett gestrichen worden. Auch die Förderung des Effizienzhauses 40 für Neubauten ist derzeit nicht verfügbar. Die Bundesregierung plant eine Neuausrichtung des Programms, um es stärker am CO₂-Einsparpotenzial auszurichten. Die Förderung des Effizienzhaus 40 für Neubauten läuft bis Ende 2022 weiter, allerdings mit geringeren Fördersätzen und einer Deckelung auf eine Milliarde Euro. Seit dem 1. März 2023 gibt es die Förderung für Neubauten nun über das Programm Klimafreundlicher Neubau (KfN), das vom BMWSB betreut wird. Diese neue Förderung, für die insgesamt 750 Millionen Euro zur Verfügung stehen, wird in Form von zinsgünstigen

Darlehen gewährt. Das KFN-Programm sieht keine Tilgungszuschüsse vor, sondern Zinsverbilligungen.

Regionale Ebene (Bundesländer)

Im Januar 2019 wurde bei der BfEE eine Koordinierungsstelle für die Energieeffizienz-Förderprogramme des Bundes und der Länder eingerichtet. Die Koordinierungsstelle pflegt eine Datenbank der relevanten Förderprogramme und intensiviert die frühzeitige Abstimmung der Förderaktivitäten von Bund und Ländern. Auf diese Weise werden die Förderstellen von Bund und Ländern dabei unterstützt, Überschneidungen und Doppelungen in der Förderung frühzeitig zu erkennen und ihre Förderprogramme besser aufeinander abzustimmen.

4.2.2. Gebäudeenergiegesetz (GEG)

Bundesministerien

Das Bundesministerium für Wohnen, Stadtentwicklung und Bauwesen (BMWSB) trägt derzeit die Gesamtverantwortung für die Umsetzung des GEG. Das Ministerium wurde im Dezember 2021 nach der Bildung der neuen Koalitionsregierung eingerichtet. Derzeit gibt es keine Ex-post-Evaluierungsverfahren für das GEG, und auf der Website des Ministeriums wurden keine Berichte veröffentlicht. Dies kann zumindest teilweise auf den typischen Zeitrahmen von 2-3 Jahren zurückgeführt werden, der für die Durchführung und Veröffentlichung formeller Evaluierungen erforderlich ist. Der derzeitige Prozess der Aktualisierung des GEG wird von BMWK und BMWSB gemeinsam koordiniert.

Bundesbehörden

Das Bundesamt für Bauwesen und Raumordnung (BBR) überwacht die Baumaßnahmen des Bundes in Berlin, Bonn und im Ausland. Das BBR unterstützt die Bundesregierung in verschiedenen Politikbereichen wie der Raumordnung, dem Städtebau, dem Wohnungswesen und dem Bauwesen. Zu seinen Aufgaben gehören unter anderem die Begleitung bedeutender Bauvorhaben, die Umsetzung von Modellvorhaben, die Baukultur und Denkmalpflege, die europäische Zusammenarbeit, die Durchführung von Architekturwettbewerben und die Durchführung von Wohnungsmarktstudien.

Das Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) ist eine Forschungseinrichtung innerhalb des BBR. Seit 2009 ist der Forschungsbereich des BBR mit dem Institut für Instandsetzung und Modernisierung von Bauwerken (IEMB) zu einer Ressortforschungseinrichtung BBSR verschmolzen. Das BBSR erbringt Forschungs- und Entwicklungsleistungen sowie wissenschaftlich fundierte Expertise für das Bundesministerium.

Regionale Ebene (Bundesländer)

Die Länder können bei der Umsetzung von Bundesgesetzen ein Veto einlegen. Die Vertreter der Länder in der zweiten Kammer (Bundesrat) müssen allen Bundesgesetzen, die ihre Finanz- und Verwaltungsangelegenheiten betreffen, zustimmen. Dies macht die Länder zu einem wichtigen Vetospieler in der Baupolitik des Bundes.

Die Länder haben ein gewisses Maß an Autonomie bei der Umsetzung der Klimapolitik. Sie können Pilotprojekte und Finanzierungsprogramme für Erneuerbare Energien, Gebäudesanierung, Wohnungsbauförderung oder Stadtentwicklung initiieren, sich in der Forschungsförderung engagieren und als Vorbild fungieren, indem sie die Energieeffizienz staatlicher Gebäude verbessern oder eine klimafreundliche Beschaffungspolitik einführen – ein beträchtlicher Ermessensspielraum für die eigene klimabezogene Baupolitik (Jacob & Kannen 2015: 5).

In Deutschland besteht ein erhebliches Vollzugsdefizit bei ordnungsrechtlichen Instrumenten. Die Durchsetzung und Umsetzung von Bundesgesetzen, einschließlich Baunormen, fällt in den Zuständigkeitsbereich der Bundesländer, denen es freisteht, die Verwaltungsstruktur ihres Vollzugssystems zu wählen [Interview 14]. Die Rechenschaftspflicht auf Bundesebene ist gering, da Datenschutzgesetze den Austausch von Unterlagen und Informationen verhindern [Interview 4]. Frühere Untersuchungen zur EnEV ergaben, dass die Nichteinhaltung der Vorschriften auf mindestens 25 % geschätzt wird, vermutlich aber noch höher liegt (Weiß und Vogelpohl, 2010: 18). Dieser mangelnde Vollzug ist eine Folge des Personalmangels in den Bundesländern (Jacob & Kannen 2015: 11).

Eine systematische Behandlung des Vollzugs auf Länderebene würde den Rahmen unserer Analyse sprengen. Dieser Bericht befasst sich mit der Bundesebene und den bestehenden Berichts- und Evaluierungsverfahren. Die Berücksichtigung dieser wichtigen Fragen bei der Umsetzung von Regulierungsmaßnahmen ist jedoch bemerkenswert, da die

Mängel im Vollzug erhebliche und negative Auswirkungen auf die Schätzungen der politischen Wirksamkeit des deutschen Policy-Mix haben würden.

4.2.3. Hilfsmittel

Bundesministerien

Das Bundesministerium der Finanzen (BMF) entscheidet über die Bereitstellung von Finanzmitteln für Klimaschutz- und Energieeffizienzmaßnahmen. Das BMF vergibt Bescheinigungen zur steuerlichen Förderung der energetischen Gebäudesanierung und erlässt eine Verordnung über die Festlegung von Mindestanforderungen an Energieeffizienzmaßnahmen in zu eigenen Wohnzwecken genutzten Gebäuden (§ 35c Einkommenssteuergesetz, Energetische Sanierungsmaßnahmenverordnung (ESanMV)). Das BMF beaufsichtigt die allgemeine Haushaltsführung (z.B. Sofortprogramm für mehr Klimaschutz 2021: 5,5 Mrd. €), das heißt es spielt eine Schlüsselrolle bei der Bewertung der fiskalischen Ausgaben und der Kosteneffizienz von Programmen.

Das BMF ist zuständig für die Vergabe von Finanzmitteln zur Förderung von Klimaschutz- und Energieeffizienzmaßnahmen. Es entscheidet maßgeblich über die Verteilung der Mittel an Initiativen, die sich mit diesen Themen befassen. So wurde im Jahr 2021 ein Sofortprogramm mit einem Budget von 5,5 Milliarden Euro eingeführt, um die Klimaschutzbemühungen zu verstärken. Mit diesem Programm sollen finanzielle Mittel für verschiedene Projekte und Initiativen zur Bekämpfung des Klimawandels und zur Förderung der Nachhaltigkeit bereitgestellt werden.

Das BMF verwaltet auch ein Zertifikat für steuerliche Anreize, das jedoch noch evaluiert werden muss. Die steuerliche Förderung zielt auf die energetische Gebäudesanierung ab. Das Zertifikat dient als Nachweis für die Einhaltung von Energieeffizienzstandards und ermöglicht es Privatpersonen oder Unternehmen, bei Gebäudesanierungen, die die vorgeschriebenen Energieeffizienzkriterien erfüllen, steuerliche Vorteile in Anspruch zu nehmen. Um Mindestanforderungen für energetische Sanierungsmaßnahmen an Wohngebäuden festzulegen, wurde die Energetische Sanierungsmaßnahmenverordnung (ESanMV) nach § 35c Einkommenssteuergesetz eingeführt. Diese Verordnung beschreibt die spezifischen Kriterien und Richtlinien, die erfüllt werden müssen, um die

Energieeffizienz von zu Wohnzwecken genutzten Gebäuden zu gewährleisten. Da es sich jedoch um eine erst kürzlich erlassene Maßnahme handelt, wurden seit ihrer Einführung keine Evaluierungen der Steuervergünstigung durchgeführt [Interview 9].

In der Vergangenheit war das Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (BMUV) für die Klimapolitik und alle Fragen im Zusammenhang mit den Erneuerbaren Energien zuständig. Das BMUV spielte eine entscheidende Rolle bei der Durchsetzung ehrgeiziger Energiestandards und der Förderung Erneuerbarer Energien, während das BMWi (jetzt BMWK) und das BMWBS weniger Neigung zu diesen Themen zeigten (Jacob und Kannen, 2015; Michaelowa, 2008; Wurzel, 2010). Das BMUV und das BMWi (jetzt BMWK) wurden oft als Organisationen mit gegensätzlichen Regulierungsvorstellungen, Interessen, kulturellen Identitäten und politischen Zugehörigkeiten angesehen (Jacob und Kannen, 2015).

Das BMUV spielt bei den derzeitigen Evaluierungsverfahren nur eine begrenzte Rolle.

Seit dem Regierungswechsel 2021 ist das BMUV nicht mehr zuständig und ist in der laufenden Legislaturperiode nicht mehr federführend an der Gesetzgebung des Gebäudeenergiegesetzes beteiligt. Die konkrete Mitwirkung des BMUV an der Ausgestaltung des Gesetzes ist deutlich zurückgegangen. Die Hauptaufgabe des BMUV in den aktuellen Evaluierungsverfahren ist die Festlegung der Emissionsfaktoren und Primärenergiefaktoren für Energieträger, die zur Hochrechnung und Berechnung der THG-Emissionseinsparungen aus Energieeinsparungen verwendet werden (Abschnitt 5.4).

Bundesbehörden

Das Umweltbundesamt (UBA) unterstützt das BMUV und koordiniert die Umweltforschung. Der Aufgabenbereich des UBA umfasst Fragen des Immissionsschutzes und des Bodenschutzes, der Abfallwirtschaft, der Wasserwirtschaft und des gesundheitlichen Umweltschutzes. Das UBA erbringt zentrale Dienstleistungen und unterstützt die Umweltforschung des BMUVs und die Koordination der Umweltforschung der Bundesbehörden.

Das UBA koordiniert den Projektionsbericht. Ziel des Projektionsberichts ist es, die konkreten Klimaschutzinstrumente zu identifizieren, mit denen die sektoralen Klimaziele des Bundes-Klimaschutzgesetzes bis 2030 erreicht werden können. Die durch den Projektionsbericht identifizierten Lücken können durch die Umsetzung effektiverer Maßnahmen

geschlossen werden. Das Szenario HIS-2030, das ähnlich wie der Projektionsbericht eine instrumentengestützte Modellierung verwendet, zeigt konkrete Handlungsalternativen auf. Die Rahmendaten, die bei der Modellierung des Projektionsberichts verwendet werden, werden vor Beginn des Modellierungsprozesses als Standardverfahren festgelegt. Die Forschungsteilnehmenden arbeiten mit dem UBA zusammen und stimmen sich mit den zuständigen Stellen ab. Um die Transparenz zu erhöhen, hat das UBA erstmals die Rahmendaten und Annahmen vor der Veröffentlichung des kommenden Berichts 2023 (UBA 2022) herausgegeben.

Öffentlich-private Einrichtungen

Die Deutsche Energie-Agentur GmbH (dena) ist eine deutsche private Einrichtung, die sich im Besitz der Bundesregierung befindet und aus mehreren Interessengruppen besteht. Sie ist als Lobbyorganisation im EU-Transparenzregister aufgeführt. Das Hauptziel der dena ist laut Satzung die Erbringung von Dienstleistungen auf nationaler und internationaler Ebene zur Gestaltung und Umsetzung der energie- und klimapolitischen Ziele der Bundesregierung, mit besonderem Fokus auf Energiewende und Klimaschutz.

Eine der Hauptaktivitäten der dena ist die Entwicklung von Pilotprojekten, um die Machbarkeit und Wirksamkeit von Energieeffizienzmaßnahmen zu testen. Diese Projekte dienen als Praxisexperimente, um die Machbarkeit verschiedener Energiespartech- niken zu bewerten. Darüber hinaus setzt sich die dena für die Verbesserung der Qualität der Planung, Umsetzung und Überwachung von Effizienzmaßnahmen ein, indem sie Standards und Richtlinien für die effektive Durchführung der Maßnahmen erarbeitet.

Die dena spielt eine Rolle bei der Bereitstellung von Informationen durch die Entwicklung von Strategien, Fahrplänen und Kommunikationsplattformen. Sie schafft und verwaltet verschiedene Kommunikationsplattformen, Strategien, Kampagnen, politische Initiativen und Veranstaltungen, um das Engagement und die Kommunikation zwischen den Beteiligten zu erleichtern. Diese Initiativen zielen darauf ab, das Bewusstsein für die Energiewende und den Klimaschutz zu schärfen und gleichzeitig die Beteiligung und Zusammenarbeit zu fördern. In Zusammenarbeit mit dem ifeu-Institut und dem Passivhaus- Institut hat die dena eine umfassende Methodik für individuelle Sanierungsfahrpläne speziell für Wohngebäude entwickelt. Der Sanierungsfahrplan dient als standardisiertes

Instrument für die Energieberatung sowohl bei Komplettsanierungen als auch bei schrittweisen Sanierungsprozessen. Er ist auf verschiedene Arten von Wohngebäuden anwendbar, darunter Einfamilienhäuser, Zweifamilienhäuser und Mehrfamilienhäuser. Der Zweck dieser Roadmap ist es, Hausbesitzende und Fachleute dabei zu unterstützen und anzuleiten, fundierte Entscheidungen zu treffen und effektive Maßnahmen für energieeffiziente Gebäudesanierungen zu ergreifen.

4.3. Berater

Evaluierungen werden an Beratungsunternehmen vergeben. Die deutsche Konfiguration von Evaluierungen und Bewertungen für die Innenpolitik, Ministerien und Bundesbehörden beauftragen eine Gruppe von Beratungsunternehmen mit der Durchführung von Ex-post-Evaluierungen und Ex-ante-Evaluierungen.

Die Auftragsvergabe erfolgt im Rahmen eines offenen Ausschreibungsverfahrens. Im Prinzip handelt es sich um offene Ausschreibungen, bei denen sich jeder bewerben kann [Interview 6]. In der Praxis gibt es eine Gruppe von Beratungsunternehmen, die in diesem Bereich tätig sind und sowohl mit der Durchführung der Ex-post-Evaluierungen als auch der Ex-ante-Evaluierungen beauftragt werden (Abschnitt 6.1 und Anhang I). Innerhalb dieses Konsortiums kommt es bei Ausschreibungen für bestimmte Projekte zu einer gewissen Variation bei den Projektleitenden und den an bestimmten Bewertungen mitwirkenden Beratungsunternehmen. Außerdem gibt es einige kleinere oder unabhängige Beratungsunternehmen. Von 2018 bis 2022 wurde z.B. die NAPE-Prognose von Fraunhofer ISI und Prognos gemeinsam durchgeführt (Fraunhofer-Leitung), ab 2023 wird der NAPE jedoch ausschließlich vom Fraunhofer ISI durchgeführt [Interview 7].

Die Evaluierungen werden von einem relativ konstanten Konsortium von Bewertenden durchgeführt. Ein Konsortium von Beratungsunternehmen (mit einigen Änderungen im Laufe der Jahre) führt kontinuierlich Bewertungen durch (Anhang I). Sie haben im Laufe der Zeit Kompetenzen und Fachwissen aufgebaut, was Pfandabhängigkeiten geführt hat sowie kumuliertem Wissen und Erfahrung im Sektor, die aufgrund des Mangels an öffentlich zugänglichen Daten von potentiellen Wettbewerbern nicht ohne Weiteres bereitgestellt werden können. Für bestimmte Themen haben einige Beratungsunternehmen

Fachwissen aufgebaut und Partnerschaften gebildet, aber es handelt sich nicht um einen hermetisch geschlossenen Kreis [Interview 6]. Diese Beratungsunternehmen folgen etablierten Praktiken und wenden in verschiedenen Sektoren die gleichen Protokolle an. Die komplexe und heterogene Natur des Bausektors stellt jedoch eine Herausforderung für die Bewertungen dar. Dies erfordert maßgeschneiderte Lösungen, um Probleme wie Rebound-Effekte zu bewältigen. Inwieweit diese Herausforderungen in der Evaluierungspraxis berücksichtigt werden, wird in Abschnitt 5.2 näher untersucht.

Die relativ kleine Gruppe von Beratungsunternehmen schränkt die Möglichkeit von Verfahrensänderungen ein. Die Bewertungs- und Beurteilungsprozesse werden in erster Linie von einer begrenzten Anzahl von Beratungsunternehmen durchgeführt. Da sowohl die Ex-post-Evaluierungen als auch die Ex-ante-Evaluierungen von derselben Gruppe von Beratern durchgeführt werden [Interview 7, 8, 9, 12], nur große und abgeschlossene Veröffentlichungen öffentlich zugänglich sind [Interview 3, 6, 7] und mehr technische Aspekte oft aus den Berichten ausgeklammert werden [Interview 7], teilweise aufgrund von Datenschutzgesetzen [Interview 6], schränkt diese vertikale Integration den Spielraum für eine Validitätsprüfung durch andere Fachleute ein.

Die Evaluierungen werden über öffentliche Ausschreibungen in Auftrag gegeben, wobei in der Auftragspezifikation die Ziele und der Umfang der Evaluierung festgelegt werden. Der in Auftrag gegebene Umfang besteht darin, die Effizienz und Wirksamkeit der zuvor durchgeführten Programme zu bewerten, und ist von Natur aus inkrementell. Es ist unwahrscheinlich, dass diese Evaluierungen zu radikalen Ergebnissen führen, die zum Beispiel das Scheitern des gesamten Programms anzeigen [Interview 12]. Es hängt vom Auftraggeber ab, was in dem öffentlichen Bericht erscheint und wie detailliert er ist [Interview 7, 12]. Oft gibt es sehr viel deutlicher formulierte Hintergrundpapiere/nicht-öffentliche Berichtsteile [Interview 7]. Die Berater im Konsortium legen Parameter und Methoden für die Durchführung von Evaluierungen innerhalb des definierten Rahmens fest [Interview 10, 11].

Die Vorgaben der Ministerien beziehen sich auf die allgemeinen Aspekte, die bewertet werden sollen, nicht aber auf Parameter oder Methoden. Die Beratungsunternehmen sind weitgehend für die Festlegung der Evaluierungsmethoden und der Parameter für die Berichterstattung über die Ergebnisse verantwortlich. Dies wird in der Regel von den

Beratungsunternehmen festgelegt, und die Vorschläge werden dann mit den Ministerien erörtert, die formell die endgültige Entscheidung treffen.

Sie sprechen mit dem Ministerium und stimmen mit ihm ab oder überprüfen es. Sie diskutieren also, aber letztendlich ist es ihre eigene Entscheidung, die Parameter festzulegen.

[Interview 11]

Die Berater haben Einfluss auf die Festlegung der Parameter. Da die Ministerien nur über begrenzte Zeit und Kapazitäten für die Validierung von Bewertungen oder Parametern verfügen, wird den Beratern eine gewisse Autonomie bei der Durchführung der Bewertungen und der Berichterstattung eingeräumt.

Aber natürlich muss ich auch sagen, dass dies, ich meine, nicht unsere Hauptarbeit hier ist. Wenn wir also eine Reihe von Beratungsfirmen haben und auch einige gute Leute, die das bewerten, dann überlassen wir ihnen auch einige Dinge. Ich meine, wenn wir es nicht selbst tun wollen oder können. Es wäre also nicht wirklich sinnvoll, auf jedes einzelne Detail einzugehen, und dann sehen wir einfach, was sie uns vorlegen, und manchmal diskutieren wir natürlich über diese Dinge und bitten vielleicht manchmal auch um Änderungen, aber im Großen und Ganzen sind wir, glaube ich, mit dem, was wir bekommen, ziemlich zufrieden.

[Interview 3].

Die Standardisierung hat in den letzten zehn Jahren zugenommen. Die Beratungsunternehmen haben sich verstärkt um eine Standardisierung der Verfahren bemüht und eine Reihe von Leitfäden erstellt [Interview, 7, 8, 9, 12], die inzwischen auch vom BMWK übernommen wurden [Interview 9]. Vor der Standardisierung der Verfahren waren die von den verschiedenen Beratungsunternehmen verwendeten Berechnungen zur Bereinigung des Endenergiebedarfs und die zur Berechnung der THG-Einsparung verwendeten Emissionsfaktoren von Beratungsunternehmen zu Beratungsunternehmen unterschiedlich und nicht transparent. Dies erschwert die Interpretation der in den vergangenen Jahren gesammelten Daten, führt zu Unsicherheiten und schränkt die Verlässlichkeit vergangener Daten ein. Dies ist besonders problematisch bei der Durchführung von Ex-ante-Arbeiten, da die Evidenzbasis, auf die zurückgegriffen werden kann, beeinträchtigt wird. Die Leitfäden enthalten eine Reihe von Verfahren zur Berechnung der Wirksamkeit und der Kosten von Maßnahmen, Instrumenten und Programmen [Interview 7, 8, 9]. Ein wichtiger Beitrag war die Festlegung einer gemeinsamen Definition von Begriffen, die

dazu beigetragen hat, eine gemeinsame Grundlage für alle Beratungsunternehmen zu erstellen [Interview 7]. In den Interviews wurde darauf hingewiesen, dass sich dies in letzter Zeit noch verstärkt hat – insbesondere im Zuge des Ukraine-Kriegs und seiner Auswirkungen auf die Energieversorgung/-preise usw. Bei der jüngsten BEG-Evaluierung (2023) wurde beispielsweise die Abweichung von den Methodikleitfäden transparent gemacht [Interview7].

Eine weitere Standardisierung von Protokollen zwischen den Beratungsunternehmen ist erforderlich. In den letzten 10 Jahren gab es zwar große Fortschritte und die Veröffentlichung des „Standardverfahrens“, aber vieles liegt noch außerhalb des Anwendungsbereichs dieser Leitfäden. Die Beratungsunternehmen wenden unterschiedliche Methoden an, um Energieanpassungen und Leistungslücken zu ermitteln [Interview 9, 12]. Die fehlende Standardisierung der Methoden bedeutet, dass die Ergebnisse nicht vergleichbar sind. Die Leitfäden sind auch in ihrem Anwendungsbereich begrenzt [Interview 12], sind relativ kurzfristig und statisch [Interview 9] und werden von den verschiedenen Beratungsunternehmen nach wie vor unterschiedlich angewendet [Interview 11]. Wir gehen auf diese Punkte in Abschnitt 5 näher ein.

5. Ex-post-Evaluierung: Verfahren, Umfang, Daten und Methoden

Dieser Abschnitt befasst sich mit Faktoren, die mit dem Umfang und der Qualität der derzeitigen Evaluierungen zusammenhängen: (i) Verfahren und Arten der bewerteten Programme; (ii) Umfang der Indikatoren; (iii) Methoden; und (iv) Daten.

5.1. Bewertungsansätze und bewertete Programme

Programme werden durch Top-down- und Bottom-up-Evaluierungsmethoden bewertet.

Top-down-Ansätze beinhalten Modellierung und statistische Extrapolation und werden zur Bewertung der Auswirkungen ökonomischer Instrumente wie Steuern und CO₂-Preisen verwendet (Abbildung 2). Bei Bottom-up-Ansätzen werden einzelne Instrumente bewertet und diese Bewertungen dann durch Extrapolation aggregiert. Diese methodischen Unterschiede bedeuten, dass der potenzielle Anwendungsbereich von Top-down- und Bottom-up-Ansätzen recht unterschiedlich ist. Wir konzentrieren uns hier auf die Bewertungsmethoden für Bottom-up-Ansätze, die eine Bewertung von ex-post gemessenen Daten erfordern.

Hintergrunddaten dienen der statistischen Extrapolation von Maßnahmen und der Bündelung von Maßnahmen und Programmen. Allgemeine „Rahmendaten“, die für die Top-down-Prozesse festgelegt werden, umfassen hauptsächlich Energiepreise (aber auch die wirtschaftliche Entwicklung), die in der Regel aus statistischen Daten des Statistischen Bundesamtes oder anderen offiziellen Datenquellen stammen [Interview 8]. Die Transparenz dieser Hintergrunddaten wird als wesentlich für die Validierung und Vergleichbarkeit der Ergebnisse bezeichnet [Interview 8]. Um die Replizierbarkeit zu erhöhen, könnten formale Datenmeldeprotokolle eingeführt werden (Abschnitt 5.4.4.).

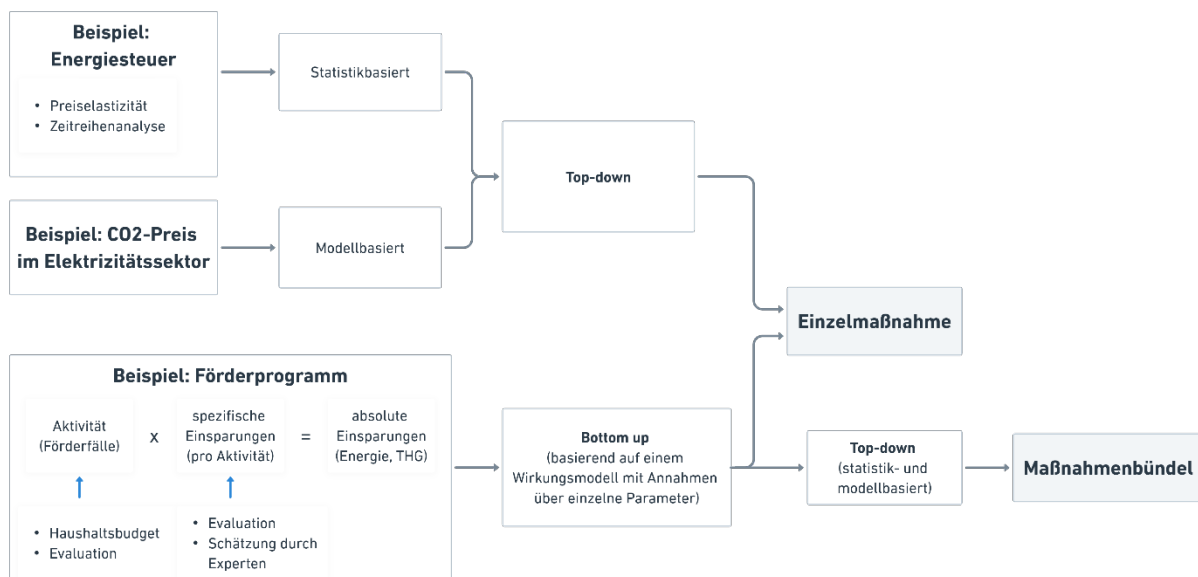


Abbildung 2: Darstellung von Bottom-up- und Top-down-Ansätzen. Angepasst von Schloman et al. (2020).

5.1.1. Steuern/wirtschaftliche Instrumente

Die Wirkungen von Energiesteueranpassungen werden durch ein Top-down-Verfahren abgeschätzt. Der CO₂-Preis (BHEG) wird als statistische Top-Down-Extrapolation und Modellierung unter Verwendung von Preiselastizitätsannahmen (KfW-EBS-WG-2018-2021, S. 57) in die Evaluierungsmethodik einbezogen (z.B. anstelle einer empirischen Schätzung von Preiselastizitäten der BEHG-Einführung). Aufgrund des begrenzten Umfangs der Evaluierungskriterien werden jedoch wichtige Aspekte, insbesondere Verteilungseffekte, nicht berücksichtigt. Dies stellt eine erhebliche Lücke in den derzeitigen Verfahren dar. Darüber hinaus schränken die relative statische Bewertung und die Nichtberücksichtigung komplexerer Systeminteraktionen, einschließlich Innovationseffekten oder Marktversagen, die Zuverlässigkeit der derzeitigen Betrachtung der Wirksamkeit von Energiesteuern und CO₂-Preisinstrumenten ein.

Auch die Datenverfügbarkeit behindert eine umfassendere Evaluierung. Die kontrafaktische Analyse wird durch Datenlücken in Bezug auf den Gebäudebestand und eine kontrafaktische Kontrollgruppe von Fällen, in denen auch ohne Fördermittel renoviert wird, eingeschränkt (KfW-EBS-WG-2018-2021, S. 61-61). Eine zeitnahe sektorale Evidenz auf dem Auflösungslevel, das für die Betrachtung von Markteffekten, sozioökonomischen

Wirkungen und Effektivität erforderlich ist, gibt es derzeit nicht. Dementsprechend werden Steuern nur in Form von Top-down-Modellen betrachtet. Die Wirksamkeit von Instrumenten zur Energie- und CO₂-Preisgestaltung im Gebäudesektor ist jedoch aufgrund der großen Heterogenität weniger vorhersehbar. Ohne die Berücksichtigung von Verteilungseffekten im derzeitigen Bewertungsrahmen gibt es zudem keinen ausreichenden Spielraum, um zu beurteilen, wie sich größere Erhöhungen der Brennstoffpreise auf die breite Bevölkerung auswirken würden:

Wir haben wirklich versucht, Daten über den Heizungsmarkt und den Baumarkt zu bekommen, als die CO₂-Steuer eingeführt wurde. Um [die Auswirkungen] von einem Monat zum anderen zu sehen. Es sind nur die Marktorganisationen, die Verbände, die diese Daten sammeln, und es gibt keine staatlichen Daten [Interview 9].

Die Einkommenssteuervergünstigungen durch das BMF für Renovierungen werden in den aktuellen Ex-post-Evaluierungen nicht berücksichtigt. Diese Form des Subventionsmechanismus ist nicht in den Evaluierungen enthalten [Interview 9]. Dieses Instrument wurde erst vor kurzem eingeführt, und die Verzögerung zwischen der Umsetzung der Maßnahmen und der Veröffentlichung der Evaluierungen könnte der Grund für diese Auslassung sein.

5.1.2. Bundesfördermittel für effiziente Gebäude (BEG)

In der Praxis wird nur das Gebäudeenergieprogramm des Bundes (BEG) nach einem Bottom-up-Ansatz evaluiert. Evaluierungen sind für das BEG-Programm aufgrund seiner direkten Kosten für den Staat erforderlich, wie vom Bundesrechnungshof (BRH) vorgeschrieben. Darüber hinaus ist die Datenerhebung eine Voraussetzung für den Erhalt von Fördermitteln aus dem Programm.

Das BEG ist das einzige politische Programm, das in die veröffentlichten Ex-post-Evaluierungsprozesse einbezogen wird. Die Ex-Post-Richtlinien konzentrieren sich weitgehend auf die Förderprogramme, die von den Bundesförderinstituten verwaltet werden. Bei den Ausgabenprogrammen des Bundes handelt es sich um die veröffentlichten Evaluierungen, die der Bottom-up-Methodik folgen (KfW-EBS-WG-2018-2021, KfW-433).

Daten über den Energieverbrauch der Begünstigten werden von den Bundesbehörden im Rahmen des Antragsverfahrens für Fördermechanismen erhoben. Informationen werden im Rahmen von Förderprogrammen als Bedingung für die Gewährung von Zuschüssen erhoben. Dabei wird der Energieverbrauch vor und nach der Installation der Maßnahmen erfasst. Darüber hinaus wird eine Umfrage bei einer Stichprobe von Fördermittelempfängern durchgeführt. Damit werden a) die Informationen aus den Anträgen validiert und b) Details zur Modellierung der Einsparungen („Gebäudebestand“) durchgeführt [Interview 7].

Es ist immer einfacher zu bewerten, wenn man Geld gibt, denn dann müssen die Leute den Fragebogen beantworten, sonst müssen sie das Geld zurückgeben. [Interview 9]

5.1.3. Gebäudeenergiegesetz (GEG)

Regulatorische Maßnahmen (Standards) sind von den derzeitigen Evaluierungsprozessen in Deutschland weitgehend ausgeschlossen. Für das GEG gibt es keine eigenen Evaluierungsprozesse oder Berichtspflichten gegenüber dem Parlament. Innerhalb des NAPE wird über das GEG nur in Bezug auf seinen Umsetzungsstatus berichtet (Abbildung 3). Der Status ist deskriptiv und enthält keine evaluativen Metriken. Im NAPE-Bericht sind unter dem GEG keine bereits durchgeführten Evaluierungen aufgeführt und es sind auch keine geplant (im Gegensatz zu anderen Instrumenten und Programmen).

Instrument	Umsetzungsstand
108. Energieeinsparrecht für Gebäude – Gebäudeenergiegesetz (GEG)	<p>Ziel: Der Deutsche Bundestag hat am 18.06.2020 das Gebäudeenergiegesetz (GEG) beschlossen, das von der Bundesregierung auf Vorschlag des Bundesministeriums für Wirtschaft und Energie und des Bundesministeriums des Innern, für Bau und Heimat eingebracht wurde. Der Bundesrat hat das Gesetz am 03.07.2020 gebilligt. Es ist am 01.11.2020 in Kraft getreten.</p> <p>Inhalt: Das Gebäudeenergiegesetz führt das Energieeinsparungsgesetz (EnEG), die Energieeinsparverordnung (EnEV) und das Erneuerbare-Energien-Wärmegesetz (EEWärmeG) zusammen und schafft ein neues, einheitliches, aufeinander abgestimmtes Regelwerk für die energetischen Anforderungen an Neubauten, an Bestandsgebäude und an den Einsatz erneuerbarer Energien zur Wärme- und Kälteversorgung von Gebäuden. Dadurch werden Anwendung und Vollzug erleichtert. Das Gebäudeenergiegesetz setzt die europäischen Vorgaben zur Gesamtenergieeffizienz von Gebäuden um und integriert die Regelung des Niedrigstenergiegebäudes in das vereinheitlichte Energieeinsparrecht. Das Gesetz behält die aktuellen energetischen Standards für Neubau und Sanierungen bei.</p>

Abbildung 3: Berichterstattung über das GEG-Programm im zuletzt veröffentlichten NAPE-Bericht. Quelle?

Das NAPE-Monitoring umfasst auch einige einzelne Regulierungsprogramme und -instrumente, aber die analytischen Inputs oder Methoden zur Berechnung der Auswirkungen sind nicht transparent. In seinem Berichtsformat werden neben Finanzierungs- und Informationsprogrammen auch Regulierungen einbezogen. In den NAPE-Berichten werden keine Ex-post-Evaluierungen für Regulierungen herangezogen und es sind auch keine geplant (BMWK 2021, S.80). Förder- und Informationsprogramme bilden die inhaltlichen Elemente der Berichte. Durch das einheitliche Berichtsformat über alle Arten von Instrumenten werden die Wirkungen von Regulierungsinstrumenten nicht ausreichend erfasst und Probleme oder notwendige Reformen innerhalb dieser Programme nicht identifiziert.

Wo es definitiv noch eine Lücke gibt, ist im Ordnungsrecht. Das Gebäudeenergiegesetz, die Energieeinsparverordnung usw., da gibt es keine Bewertung. Die Abschätzung, wie sich eine Verschärfung des Ordnungsrechts langfristig auf die Emissionen auswirkt, ist also alles sehr, sehr modellbasiert und annahmegetrieben. Es gibt einfach keine Daten dazu. [Interview 6]

Derzeit gibt es keine spezifischen Anforderungen für die Bewertung des Programms GEG. Dies ist möglicherweise darauf zurückzuführen, dass es der Regierung keine direkten Kosten verursacht und in der Vergangenheit nicht im gleichen Maße wie die Steuer Ausgaben geprüft wurde, obwohl es erhebliche Governance-Anforderungen und Verwaltungskosten verursacht [Interview 13]. Dies könnte auch auf die Durchführungsbestimmungen auf Länderebene und eine lokale vertikale Koordination mit den Bundesministerien zurückzuführen sein [Interview 4]. Infolgedessen hat sich der Schwerpunkt der Evaluierung in erster Linie auf das BEG-Programm konzentriert, während Evaluierungen für das GEG-Programm nur begrenzt oder gar nicht vorhanden sind. Es gibt andere in Auftrag gegebene Studien, die sich mit spezifischen Themen wie der Leistungslücke bei der Gebäudesanierung befassen (siehe z.B. Jagnow und Wolff, 2020). Diese sind jedoch nicht Teil der formalen Evaluierungsverfahren.

5.1.4. Methodikleitfäden

Die aktuellen Leitfäden für die Bewertung scheinen speziell auf die Bewertung von Förderprogrammen und Zuschüssen zugeschnitten zu sein. Theoretisch sind die Leitfäden auch auf andere Arten von Programmen oder Instrumenten anwendbar und enthalten eine Evaluierungslogik, die diese einbeziehen kann (Ex-post-Leitfaden, S. 28). Der praktische Leitfaden für Indikatoren hat jedoch einen klaren Schwerpunkt auf der Bewertung von Subventionen. Regulierungsmaßnahmen werden zwar in die Interventionslogik einbezogen, erhalten jedoch nur begrenzte Anleitungen und verfügen im Vergleich zu anderen Maßnahmenarten über die geringste Anzahl geeigneter Indikatoren (Ex-post-Leitfaden, S. 49-51). Darüber hinaus wird die Verwendung von Effizienzindikatoren für Regulierungsinstrumente nicht empfohlen, da sie nur begrenzt vergleichbar sind und die Effizienz der Finanzierung möglicherweise falsch wiedergeben (Ex-post-Leitfaden, S. 62). Aus den Interviews geht hervor, dass die Leitfäden den damaligen Entwicklungsschwerpunkt widerspiegeln und der Methodikleitfaden als „work in progress“ zu betrachten ist [Interview 7].

In den Leitfäden wird zwischen verschiedenen „Interventionslogiken“ unterschieden. Regulierungsmaßnahmen und Steuern werden als eigene Logiken zu den Subventionsprogrammen betrachtet (Ex-post-Leitfaden, S. 29). Bei Maßnahmen mit ordnungspolitischer Interventionslogik geht es in der Regel um die Festlegung verbindlicher Regeln in Gesetzen oder Verordnungen (legislative Maßnahmen). Damit werden Verpflichtungen (z.B. zu Verhaltensweisen, Marktzugang) oder Standards definiert, deren Nichteinhaltung auf bestimmte Weise sanktioniert wird. Praktische Beispiele für Maßnahmen mit ordnungsrechtlicher Interventionslogik sind: ordnungsrechtliche Maßnahmen (z.B. EnEV-Anforderungen für Gebäude) und Steuern oder Abgaben (z.B. KWKG).

Die Bundeshaushaltsordnung (BHO) unterscheidet zwischen ordnungsrechtlichen und fiskalischen Instrumenten. Den Rahmen für die weitere Ausgestaltung der Maßnahme, d.h. die Umsetzung der politischen Vorgabe, geben die Bundeshaushaltsordnung (BHO) und ihre Verwaltungsvorschriften vor. Im Wesentlichen werden zwei Arten von Maßnahmen unterschieden: zum einen gesetzgeberische Ordnungsmaßnahmen und zum anderen finanzwirksame Maßnahmen. Aufgrund der Thematik der BHO erfolgt diese Unterscheidung aus haushaltsrechtlicher Sicht/Logik, d.h. es wird unterschieden, ob für die

Durchführung der Maßnahme Haushaltsmittel bereitgestellt werden oder nicht (Ex-post-Leitfaden, S. 24).

Es besteht eine Diskrepanz zwischen dem Bewertungsschwerpunkt der BHO und den Methodikleitfäden. Die BHO legt eine inputorientierte Bewertung fest (wie viel müssen wir ausgeben, um ein bestimmtes Ergebnis zu erzielen), während in den Leitfäden ein outputorientierter Ansatz vorgeschlagen wird, der die Wirkung von Subventionen und Vorschriften besser vergleichbar macht als einfache Effizienzmessungen:

Der Vorteil der wirkungsorientierten Evaluationsperspektive liegt darin, dass die Vergleichbarkeit verschiedener Maßnahmentypen durch die Standardisierung von Wirkungsmodellen und deren Interventionslogiken ermöglicht wird. So können Maßnahmenpakete bedarfsgerecht zusammengestellt und Förderangebote systematisch vorbereitet werden, oder es können potenziell auftretende Überschneidungseffekte identifiziert und individuell berücksichtigt werden (z.B. bei der „Übergabe“ verschiedener Maßnahmen in den einzelnen Schritten des Wirkungsmodells, wie z.B. die geförderte Investition im Anschluss an eine geförderte Beratung) (ex-post-Leitfaden, S. 32).

Der Leitfaden rät davon ab, Effizienzindikatoren für Regulierungsinstrumente zu verwenden. Der Leitfaden legt fest, dass bei der Interpretation des Indikators „Fördereffizienz“ auch die Art der Interventionslogik, die mit einer Maßnahme verbunden ist, berücksichtigt werden sollte (Ex-post-Leitfaden, S. 62). Konkret heißt es im Leitfaden, dass Maßnahmen mit regulatorischer Interventionslogik im Vergleich zu finanziell wirksamen Maßnahmen mit ökonomischer oder informatorischer Interventionslogik zunächst eine hohe Fördereffizienz aufweisen können. Diese Wahrnehmung beruht jedoch in erster Linie darauf, dass ordnungspolitische Maßnahmen keine staatlichen Finanzhilfen vorsehen. Es ist wichtig zu beachten, dass bei der Berechnung des Indikators für die Fördereffizienz andere öffentliche und private Kosten, die mit Regulierungsmaßnahmen verbunden sind, wie z.B. Verwaltungskosten für die Kontrolle oder Kosten für Personal und Investitionen, die zur Einhaltung bestimmter Grenzwerte oder Mindeststandards erforderlich sind, oft nicht berücksichtigt werden. Daher kann eine Regulierungsmaßnahme aus wirtschaftlicher Sicht insgesamt eine deutlich geringere Fördereffizienz aufweisen, als der Indikator ursprünglich anzeigt. Bei Maßnahmen mit regulatorischer Interventionslogik wird emp-

fohlen, mit den Vertragspartnern zu klären, ob der Indikator für die Fördereffizienz überhaupt ausgewiesen werden sollte. Diese Entscheidung sollte auf der Möglichkeit beruhen, relevante Kostenfaktoren in die Berechnung des Indikators einzubeziehen. Wenn die Entscheidung getroffen wird, den Indikator auszuweisen, ist es notwendig, eine klare Klassifizierung vorzunehmen und einen Hinweis auf die begrenzte Vergleichbarkeit der Fördereffizienz mit finanziell effektiven Maßnahmen zu geben.

Der derzeitige Evaluierungsrahmen ist für die Bewertung von regulatorischen Maßnahmen ungeeignet. Die Kosten, die durch Regulierungen entstehen, sind makroökonomischer und administrativer Natur und werden im derzeitigen Evaluationsrahmen nicht ausreichend berücksichtigt. Die derzeitigen Evaluierungsprozesse dienen in erster Linie als Verfahren zum Nachweis der Kosteneffizienz von Förderprogrammen des Bundes (Abschnitt 6.2). Da Regulierungskosten nicht als direkte Kosten des Staates betrachtet werden, wurde diesem Politikbereich nicht die gleiche Aufmerksamkeit geschenkt wie den Steuerausgaben, und er unterliegt nicht der Kontrolle des BRH. Dabei werden die beträchtlichen potenziellen makroökonomischen und wohlfahrtsfördernden Auswirkungen der Einführung von Vorschriften und die Verwaltungskosten für eine wirksame Verwaltung und glaubwürdige Durchsetzung dieser Vorschriften zur Gewährleistung ihrer Wirksamkeit übersehen. Die Berichtspflichten im Rahmen des KSG sind relativ neu und die Evaluierungsverfahren wurden nicht an die notwendigen Kriterien für eine wirksame Evaluierung von Regulierungsmaßnahmen angepasst.

5.2. Umfang der Bewertungsindikatoren

In diesem Abschnitt geht es um den Umfang der Ex-post-Evaluierungen unter Verwendung eines erweiterten Kriterienkatalogs als normative Benchmark (Tabelle 7).

Dimension	Komponenten	Ex-post-Leitfaden zur Methodik	Veröffentlichte Evaluierungen
Effektivität	<i>Energienutzung/ Einsparungen</i>	Endenergieeinsparung (Lebensdauer) <ul style="list-style-type: none"> Gebäudemodell (energetische Gebäudebilanzierung). 	Wird in allen zugänglichen Evaluierungen berechnet.
	<i>THG-Einsparung</i>	Treibhausgasreduzierung (Lebensdauer) <ul style="list-style-type: none"> Basierend auf dem Emissionsfaktor. Statisch. 	Wird in allen zugänglichen Evaluierungen berechnet.
	<i>Zusammensetzung des Bestands</i>	Anzahl der Heizungsanlagen <ul style="list-style-type: none"> Anzahl der Heizungsanlagen als Anteil am Gesamtbestand. 	In zugänglichen Evaluierungen nicht enthalten.
	<i>Interaktionseffekte</i>	Synergien zwischen politischen Maßnahmen auf Bündelungsebene.	Synergien werden auf der Grundlage der Umfrageergebnisse bewertet, haben aber keinen Einfluss auf die endgültigen Einsparungen.
	<i>Rebound</i>	Nicht enthalten <ul style="list-style-type: none"> Zu schwierig zu berechnen (S. 94) 	In zugänglichen Bewertungen nicht enthalten.
Kosten-effizienz	<i>Dynamische Kosteneffizienz</i>	Auswirkungen der Investitionen <ul style="list-style-type: none"> Auf der Grundlage einer Expertenbewertung Verursachte Investitionen (S. 66) Hebeleffekt (S. 67) Geschaffene/gerettete Arbeitsplätze (S. 67-68) 	Ist in allen Evaluierungen enthalten außer denen zur Evaluation von Kosteneffizienz. Keine der für diesen Report kodierten Evaluierungen berechnet sie.
	<i>Statischer Wirkungsgrad</i>	Marginale Vermeidungskosten (S. 65) Grenzkosten der Energieeinsparung (S. 63-64)	Wird in allen zugänglichen Evaluierungen berechnet.
Steuerliche Belastung	<i>Kosten/Einnahmen für den Staat</i>	Gesamtkosten des Steuerhaushalts	Wird in allen zugänglichen Evaluierungen berechnet.
Vertrieb	<i>Auswirkungen auf die Bevölkerung</i>	Sozioökonomische Auswirkungen <ul style="list-style-type: none"> Nicht enthalten. 	Die jüngste BEG-Evaluierung umfasst „soziale Aspekte“, die Zielgruppen nach Einkommen, Alter und Bildung umfassen.
		Schaffung von Arbeitsplätzen <ul style="list-style-type: none"> Angewandter Investitionsfaktor. Bewertung durch Experten. 	Wird in allen zugänglichen Evaluierungen berechnet.
		Zielgruppen/Empfänger der Mittel <ul style="list-style-type: none"> Die Zielgruppendefinition ist ein entscheidendes Element des Zielsystems (S. 41-42). 	<ul style="list-style-type: none"> KfW 433 zielt speziell auf „Early Adopters“ ab In anderen Evaluierungen werden die Gruppenmerkmale der Begünstigten untersucht, ohne dass eine ausdrückliche Ausrichtung auf bestimmte Zielgruppen erfolgt. Die BEG-Evaluierung umfasst „soziale Aspekte“, die Zielgruppen nach Einkommen einschließen.
		Auswirkungen auf die Heizkosten <ul style="list-style-type: none"> Senkung der Energiekosten (S. 65-66) 	Berechnet in der BEG-Evaluierung von 2023.
	<i>Auswirkungen auf die Unternehmen</i>	Schaffung von Arbeitsplätzen	Wird in allen zugänglichen Evaluierungen berechnet.

Dimension	Komponenten	Ex-post-Leitfaden zur Methodik	Veröffentlichte Evaluierungen
		Wettbewerbsfähigkeit • Nicht enthalten.	In zugänglichen Evaluierungen nicht enthalten.
Akzeptanz	<i>Bevölkerung</i>	Zufriedenheit mit den Maßnahmen	Wird in allen Evaluierungen außer MAP ermittelt.
	<i>Firmen</i>	Zufriedenheit der Begünstigten mit dem Verwaltungsverfahren	Wird in allen Evaluierungen außer MAP ermittelt.
	<i>Politisch</i>	Nicht enthalten	Nicht enthalten.
Governance	<i>Verwaltung</i>	Verwaltungskosten • Kosten der Bundesbehörde für die Verwaltung des BEG.	In zugänglichen Evaluierungen nicht enthalten.
		Verwaltung der Maßnahmen • Zufriedenheit der Begünstigten mit dem Verwaltungsverfahren.	Wird in allen Auswertungen außer MAP berechnet.
	<i>Anforderungen an Informationen</i>	Qualität der Daten	Veröffentlichte Evaluierungen geben ihre Datenquellen an. Die Qualität wird nicht explizit überprüft.

Tabelle 7: Umfang der Ex-post-Evaluierungen. Die Methodikleitfäden werden zusammen mit der Anwendung in den veröffentlichten Bewertungen dargestellt, um einen Vergleich mit ihrer Anwendung in der Praxis zu ermöglichen. Die gelbe Farbe zeigt an, dass der Indikator teilweise enthalten ist, während die rote Farbe bedeutet, dass er nicht enthalten ist.

5.2.1. Wirksamkeit

Bei den Bewertungen wird die Wirksamkeit der Maßnahmen in erster Linie durch die Ermittlung des Energiebedarfs vor und nach der Umsetzung der Maßnahmen beurteilt.

Allerdings gibt es Herausforderungen im Evaluierungsprozess, wie z.B. die fehlende Standardisierung bei der Berechnung des Energiebedarfs und die mangelnde Klarheit hinsichtlich der Sicherheit und Unsicherheit dieser Berechnungen (Abschnitt 5.4).

Die wichtigsten Indikatoren für die Wirksamkeit der durchgeführten Maßnahmen sind Energieeinsparungen und die Verringerung von Treibhausgasen (Ex-post-Leitfaden, S. 50-51). Der Evaluierungsprozess umfasst eine Kombination von Bottom-up-Evaluierungen einzelner Maßnahmen, die dann zur Ermittlung der Einsparungen auf Programmebene aggregiert werden.

THG-Einsparung

Primärenergieeinsparungen werden zur Berechnung der THG-Einsparung verwendet.

Es scheint, dass die Primärenergiefaktoren statisch angewendet werden und besonders empfindlich auf den Primärenergiefaktor für Strom reagieren, der regelmäßig aktualisiert werden muss (Ex-post-Leitfaden, S. 18). Die Befragten wiesen auch darauf hin, dass die Berater in der Praxis unterschiedliche Werte für diese Faktoren anwenden können.

„Wir nehmen auch andere Kohlenstoffemissionsfaktoren, die wir für präziser halten“. [Interview 11]

Die Berechnung der THG-Einsparung erfolgt für die geschätzte Lebensdauer der installierten Maßnahmen. Die verwendeten Primärenergiefaktoren sollten jedoch mit dem Auftraggeber abgestimmt werden, um veränderte CO₂- und Primärenergiefaktoren zu berücksichtigen. Eine aktuelle Prognose dieser Faktoren sollte zum Zeitpunkt der Bewertung verwendet werden (Ex-post-Leitfaden, S. 19-20). Allerdings gibt es wenig Anleitung, wie diese Prognosen umzusetzen sind und daher in der Praxis unterschiedlich angewendet werden können.

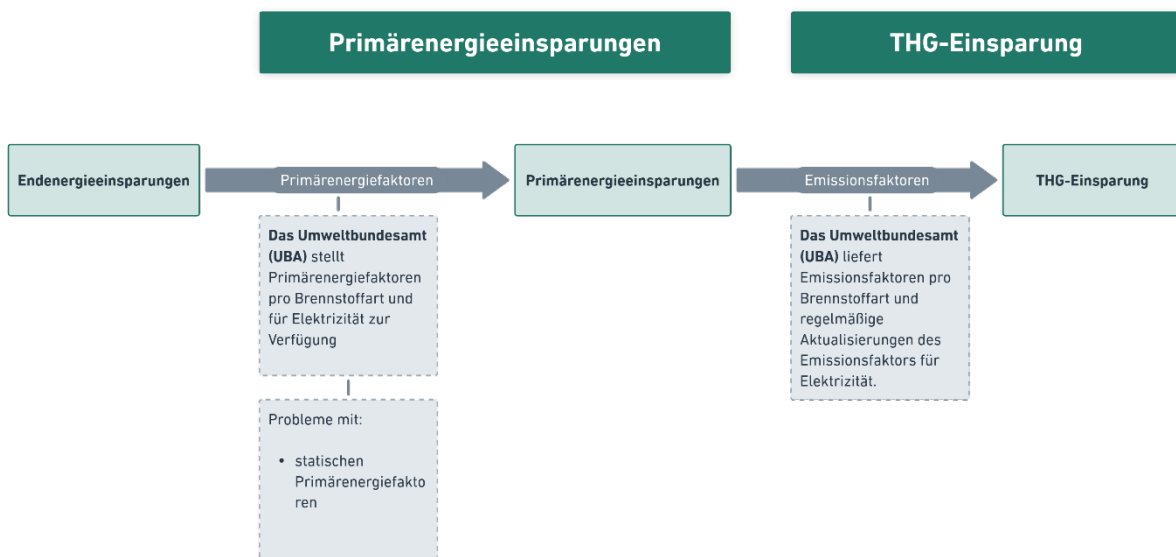


Abbildung 4: Metrische Anpassungen, die auf die Endenergieeinsparungen angewendet werden, um die THG-Einsparung in den Methodikleitfäden zu bestimmen. Quelle: Eigene Darstellung

Die Anwendung von Emissionsfaktoren ist in den veröffentlichten Bewertungen nicht transparent. In den analysierten Evaluierungen wird nicht transparent beschrieben, wie die Emissionsfaktoren für die Vorhersage künftiger Energieeinspareffekte angewendet wurden. Nach unserer Einschätzung scheinen die Emissionsfaktoren statisch angewandt worden zu sein (d.h. ein aktualisierter Referenzwert zum Zeitpunkt der Bewertung), während die Kosteneinsparungen durch installierte Maßnahmen im Zeitverlauf geschätzt werden. Die Energieeinsparungen basieren auf dem erwarteten Lebenszyklus der Maßnahme, aber die statische Anwendung führt z.B. zu einer Unterschätzung der THG-Einsparung durch die Installation von Wärmepumpen. Da erwartet wird, dass die Stromversorgung in den nächsten 15-20 Jahren dekarbonisiert wird, würde die erwartete jährliche

CO₂-Einsparung über den Lebenszyklus der Maßnahme steigen. Lässt man dies außer Acht, erscheinen Wärmepumpen in Bezug auf die THG-Einsparung potenziell weniger effizient. Umgekehrt verhält es sich bei Renovierungen. Renovierungen bieten unter den derzeitigen Bedingungen das größte Potenzial zur Verringerung der Treibhausgasemissionen, wenn sie mit einem Gaskessel kombiniert werden, oder wenn der Strommix kohlenstoffintensiver ist (wenn er mit einer Wärmepumpe gekoppelt wird). Diese Berechnung sollte mit potenziellen Konfidenzintervallen berechnet werden, da sie in hohem Maße von der Geschwindigkeit der Dekarbonisierung der Stromversorgung abhängt. Es wird daran gearbeitet, dies in der laufenden BEG-Evaluierung zu ändern [Interview 7].

Rückstoß

In den Leitfäden wird empfohlen, bei Ex-post- oder Ex-ante-Evaluierungen keine Rebound-Effekte zu berechnen. Rebound-Effekte sind gegenläufige Wirkungen einer Maßnahme. Sie umfassen direkte und indirekte Effekte. Direkte Rebound-Effekte, wie z.B. die energiesparende LED-Technologie, führen zu einem höheren Verbrauch innerhalb desselben Systems. Indirekte Rebound-Effekte führen zu einem erhöhten Verbrauch in einem System aufgrund von Einsparungen in einem anderen System, wie z.B. mehr Fahrten mit einem kraftstoffsparenden Auto. Die Abgrenzung von Rebound-Effekten von anderen Verhaltenseffekten gilt als schwierig und ist Gegenstand anhaltender wissenschaftlicher Debatten. Daher wird in den Leitfäden davon abgeraten, sie einzubeziehen, vor allem wegen der Komplexität und des Mangels an robusten Methoden zur Quantifizierung dieser Effekte, wodurch ihre Messung und Einbeziehung oft spekulativ und weniger zuverlässig ist (Ex-post-Leitfaden, S. 94). In der laufenden BEG-Evaluation wird geprüft, wie der Verbrauchsvergleich nach dem Institut für Wohnen und Umwelt (IWU), d.h. ein empirisch analysierter Einfluss des individuellen Verhaltens, in die Evaluation einbezogen werden kann. Dies wurde in dem kürzlich veröffentlichten Bericht berücksichtigt. Dies ist zwar nicht direkt der Rebound, wird aber teilweise indirekt berücksichtigt [Interview 7].

5.2.2. Kosten/Kostenwirksamkeit

Die Kosteneffizienz ergibt sich aus den berechneten Energieeinsparungen und den fiskalischen Kosten des Programms. Die Kosteneffizienz von Energieeffizienzprogrammen

wird durch die Betrachtung der erzielten Endenergieeinsparungen und der damit verbundenen fiskalischen Programmkosten bestimmt (Ex-post-Leitfaden, S. 60-61). Dies ist eine relativ statische Perspektive auf die Wirksamkeit. Eine dynamischere Perspektive auf die Kosteneffizienz könnte Projektionen für die zukünftigen Kosten und den Bedarf an Strom und Energie beinhalten, die mit der pro Maßnahme eingesparten Energie verglichen werden [Interview 7, 9]. Dabei könnten auch Kostenverläufe für emissionsarme/-freie Technologien berücksichtigt werden, einschließlich potenziell einschränkender Faktoren wie Lieferketten, Verfügbarkeit von Materialien, Fertigkeiten und Fähigkeiten zur Installation von Maßnahmen.

Die Berechnung der Reduzierung von Treibhausgasen (THG) beruht auf der Anwendung von Primäremissionsfaktoren. Die derzeitige Methodik scheint aufgrund der Anwendung von Energiefaktoren die Förderung von Renovierungen zu begünstigen (siehe oben: THG-Einsparung). Infolgedessen können die Grenzkosten der Emissionsminderung für kostenintensivere Renovierungen niedriger erscheinen, als sie tatsächlich sind, wenn man sie dynamisch über die Zeit betrachtet. Dieser Vorbehalt weist auf eine Einschränkung des derzeitigen Evaluierungsverfahrens hin, da es Renovierungen unterstützen kann, die bei Berücksichtigung künftiger Kosten nicht als kosteneffizient gelten würden. Er unterstreicht die Notwendigkeit, die langfristigen Kosten und Vorteile von Energieeffizienzmaßnahmen zu bewerten und zu berücksichtigen, um eine umfassende Evaluierung ihrer Kosteneffizienz zu gewährleisten.

Ich möchte auch auf eine Gefahr von Evaluierungen hinweisen. Ich meine, man neigt dazu, Indikatoren zu definieren, und einer der wichtigsten Indikatoren, nach dem die Politik immer fragt, ist die Treibhausgaseffizienz (eingesparte Tonnen pro Euro). Ich denke, das führt in kürzester Zeit in eine Sackgasse, denn wenn wir unseren Strom und unsere Fernwärme und sogar das Gas usw. dekarbonisieren, werden wir in Zukunft keine Treibhausgase mehr einsparen, weder mit effizienteren Geräten noch mit Gebäuden noch mit anderen Energie-sparmaßnahmen noch mit anderen Modellwechselmaßnahmen. Denn das ist alles dekarbonisiert, also man spart erneuerbare Energie. Und ich glaube, es gibt eine sehr mechanische Sichtweise. Es war die gleiche Diskussion, die wir bei der Evaluierung des Marktanzreizprogramms hatten [...] Die Evaluierung geht also nicht wirklich darauf ein, inwieweit

das einzelne Thema, das finanziert, unterstützt oder reguliert wird, wirklich für das gesamte Energiesystem notwendig ist. Man muss sich also immer der Grenzen einer Evaluierung und der verwendeten Indikatoren bewusst sein, insbesondere dieser Koeffizienten, Euro pro Tonne, denn sowohl im Zähler als auch im Nenner besteht Unsicherheit. Es ist also sehr instabil und führt zu falschen Schlussfolgerungen, wenn man sich nur an diese Vorgaben hält. [Interview 9].

Dynamische Effekte werden unterschätzt, einschließlich möglicher Markt- und Systemfehler. Bei den derzeitigen Evaluierungsverfahren werden die dynamischen Aspekte der Kostenwirksamkeit sowie das Verhalten der Marktakteure und des Einzelnen meist übersehen. Faktoren wie Kurzsichtigkeit, begrenzte Rationalität oder Marktzutrittsschranken, die das Verhalten der Akteure/Verbrauchenden beeinflussen können, sind unterentwickelt. Obwohl die Forschung begonnen hat, diese Aspekte zu erforschen (z.B. IREES), ist die Integration in die Evaluierungsprozesse begrenzt. Einige Überlegungen zu dynamischen Effekten werden durch die Definition von „Effektbereinigung“ (Abschnitt 5.4.2.) berücksichtigt, aber diese werden in veröffentlichten Evaluierungen aufgrund mangelnder Qualitätsdaten in der Regel nicht angewandt. Das Verständnis und die Einbeziehung dieser verhaltensbezogenen Faktoren in die Evaluierung kann wertvolle Erkenntnisse über die Wirksamkeit von Maßnahmen und Programmen liefern [Interview 5].

Warum treffen die Menschen keine rationalen Entscheidungen? Welche Hindernisse gibt es? Natürlich, wenn man sich die einzelnen Modelle anschaut, haben sie schon unendlich viele Erkenntnisse aus dem Konglomerat der Analysen. Aber ich denke, es gibt noch viel zu erforschen. [...] Ich habe mich in letzter Zeit mehr für die Entscheidungsmodelle der einzelnen Akteure interessiert [Interview 5].

Eine begrenzte dynamische Perspektive hat möglicherweise sehr hohe Renovierungsstandards mit hohen Grenzkosten gegenüber der Ausrichtung auf einen größeren Wohnungsbestand begünstigt. Die derzeitige Verfahrenslogik begünstigt den Fokus auf ein einzelnes Haus bei Renovierungen. Dies schafft Anreize für hohe Effizienzstandards in einer einzelnen Wohneinheit [Interview 14]. Die Fördermittel für KfW-Effizienzstandards bezuschussen einen höheren Anteil der Kosten für höhere Effizienzstandards (Tabelle 8). Der maximale Zuschuss für ein „Effizienzhaus40“ beträgt 24.000 Euro, in Verbindung mit einer Erneuerbare-Energien-Anlage bis zu 37.000 Euro. Wärmepumpen werden im

Vergleich dazu mit maximal 40 % gefördert (BAFA 2023). Da die Anschaffungskosten niedriger sind (10.000-30.000 Euro), entspricht dies einem maximalen Zuschuss von 12.000 Euro, der sich mit sinkenden Kosten durch Innovation und Skaleneffekte verringert. Die Grenzkosten für die durch die Renovierung eingesparte zusätzliche Energie sind relativ hoch. Eine alternative Logik würde darauf abzielen, eine größere Anzahl von Häusern auf ein angemessenes Effizienzniveau zu bringen, damit eine Wärmepumpe effektiv funktioniert. Es wurde vorgeschlagen, dass dieser Wert bei etwa KfW 70 liegt [Interview 14], aber es sind weitere Untersuchungen erforderlich, um einen zuverlässigen Richtwert zu ermitteln. Wenn dies zutrifft, könnte mit demselben Steuerbudget potenziell die doppelte Anzahl von Wohneinheiten renoviert und die Installation von Wärmepumpen in größerem Umfang beschleunigt werden. Bei einer dynamischen Betrachtung könnte der derzeitige Ansatz eine geringere Kostenwirksamkeit aufweisen als bei einer statischen Betrachtung, insbesondere wenn die Stromversorgung dekarbonisiert wird.

KfW-Standard	Zuschuss in % der Gesamtkosten (pro Wohneinheit)	Insgesamt (Euro pro Wohneinheit)
Effizienzhaus 40 + erneuerbare Energie	25 % von max. 150.000 EUR	37,500
Effizienzhaus 40	20 % von max. 120.000 EUR	24,000
Effizienzhaus 55 + erneuerbare Energie	20 % von max. 150.000 EUR	30,000
Effizienzhaus 55	15 % von max. 120.000 EUR	18,000
Effizienzhaus 70 + erneuerbare Energie	15 % von max. 150.000 EUR	22,500
Effizienzhaus 70	10 % von max. 120.000 EUR	12,000
Effizienzhaus 85 + erneuerbare Energie	10 % von max. 150.000 EUR	15,000
Effizienzhaus 85	5 % von max. 120.000 EUR	6,000

Tabelle 8: Verfügbare Fördermittel für KfW-Standards. Quelle: KfW (2023).

Eine dynamischere Perspektive auf die Kosteneffizienz muss die erwarteten Kostenverläufe für die Dekarbonisierung der Stromerzeugung mit der Reduzierung der Nachfrage (Energieeinsparungen) kombinieren. Um zu bewerten, ob die Energieeinsparungen bei höheren Renovierungsstandards kosteneffektiv sind, müssen die Projektionen für die Kosten der Dekarbonisierung der Stromerzeugung pro kWh und die voraussichtlichen Kostenverläufe für die Einführung von Wärmepumpen integriert werden. Wärmepumpen sind nachweislich die kosteneffizienteste Methode zur Reduzierung des deutschen Gasverbrauchs (Altermatt et al., 2023), und um die Dekarbonisierungsziele zu erreichen, muss auch die Stromerzeugung dekarbonisiert werden. Modellierungen zeigen, dass ein

ausgewogener Ansatz zwischen Nachfragereduzierung und Dekarbonisierung des Angebots erforderlich ist [Interview 9]. Dies erfordert einen stärker integrierten strategischen Ansatz, der die Sektorenkopplung und die Abschätzung der Nachfrage im Energiesystem kombiniert. Das Fehlen einer umfassenden Behandlung von Interaktionseffekten bei der Bewertung von Programmen (Abschnitt 5.4.3.) deutet auf einen potenziellen Mangel an Weitsicht hin und zeigt einen Bereich auf, in dem mehr methodische Aufmerksamkeit erforderlich ist.

Makroökonomische Indikatoren

Senkung der Energiekosten. Dieser Indikator bewertet das Ziel der Energiekostensenkung und ist Teil der Zielerreichungs- (brutto) und Wirkungsanalyse (netto einschließlich Anpassung) (Ex-post-Leitfaden, S. 65-66). Die Energiekosteneinsparungen werden auf der Grundlage der berechneten Energieeinsparungen über den gesamten Wirkungszeitraum der Maßnahme berechnet. Die Senkung der Energiekosten wird in der Stichprobe der koordinierten Ex-post-Evaluierungen in der kürzlich veröffentlichten Bewertung des BEG berücksichtigt.

Ausgelöste Investitionen und Hebeleffekt. Bei einer wirtschaftlichen Betrachtung spielen Wertschöpfungseffekte durch die von der Maßnahme ausgelösten Investitionen eine wesentliche Rolle. Die Hebelwirkung ist ein Verhältnis zwischen den eingesetzten Ressourcen und den ausgelösten Investitionen (Ex-post-Leitfaden, S. 65-66). Bei diesem Indikator bleiben die Verwaltungskosten unberücksichtigt. Der Hebeleffekt gibt an, wie viel Investitionen pro eingesetztem Euro ausgelöst werden. In der Praxis werden sowohl indirekt als auch direkt ausgelöste Investitionen berechnet. Die KfW-433-Evaluierung hebt im Vergleich vor allem die Hebelwirkung der Förderung hervor und ermittelt, inwieweit sie zusätzliche private Investitionen auslöst (KfW-433, S. 101-102). In den Evaluierungen KfW-EBS-WG-2018-2021 (S. 9-10) und MAP 2019 (S. 27-28) werden dagegen vorrangig Input-Output-Modelle zur Berechnung der direkten bzw. indirekten Wertschöpfungseffekte herangezogen. Dabei werden die Investitionen je nach Art den verschiedenen Sektoren zugeordnet und die entsprechenden Nachfrage- und Produktionssteigerungen innerhalb dieser Sektoren analysiert. Bei den Auswertungen der nationalen Monitoring-Berichte und der Bewertung der Hocheffizienzpumpen werden diese Effekte nicht berücksichtigt.

Wir versuchen immer, das ein wenig zu tun, aber natürlich ist die Zeit begrenzt, so dass man das nicht hundertprozentig durchziehen kann. Andererseits werden diese makroökonomischen Auswirkungen und dergleichen normalerweise in separaten Projekten untersucht. Sie sind also nicht wirklich Teil der formalen Bewertung. Aber wir hatten zum Beispiel 2010 ein Projekt über die makroökonomischen Auswirkungen verschiedener Effizienzmaßnahmen und untersuchten die verschiedenen Auswirkungen auf Arbeitsplätze, Energiesicherheit usw. Dies wird also in der Regel in separaten Aktivitäten durchgeführt, was man von der eigentlichen Bewertung des politischen Instruments nicht wirklich erwartet.
[Interview 9]

Die Einbeziehung verbesserter Methoden könnte ausdrücklich in die Ausschreibungsspezifikation aufgenommen werden. Die Modellierung erfordert Ressourcen des Bewertenden, und wenn der Auftraggeber diese Frage nicht ausdrücklich stellt („Wie hoch ist der mit der Finanzierung verbundene Bruttowertschöpfungseffekt?“), wird sie nicht einbezogen. Andernfalls würden die zusätzlich anfallenden Kosten den Prozess verteuern – was sich im Wettbewerb um den Auftrag möglicherweise negativ auswirkt und dazu führen könnte, dass der Auftrag an einen Mitbewerber vergeben wird. Aus Sicht der Gutachter ist es daher wirtschaftlich sinnvoll, nur den in der Leistungsbeschreibung genannten Leistungsumfang anzubieten [Interview 7].

Die Schaffung von Arbeitsplätzen wird berücksichtigt und ist in der Evaluierungsmethodik als Skalenfaktor enthalten. Die Art und Weise, wie dieser Faktor angewandt wird, widerspricht eigentlich sozioökonomischen Überlegungen, denn die Indizes begünstigen teurere Maßnahmen, die einen höheren Index für die Schaffung von Arbeitsplätzen aufweisen. Dies begünstigt daher teurere Renovierungen, die derzeit von der KfW gefördert werden, gegenüber kostengünstigeren Renovierungen, die eher einkommensschwachen Haushalten zugutekommen.³ Dies kann zu einer Verzerrung der Kosteneffizienz der Maßnahmen führen, was potenziell der Fall ist. Eine zweite Überlegung ist, dass Deutschland derzeit mit einem Arbeitskräfte- und Fachkräftemangel bei der Durchführung der Maß-

³ Die aktuellen Fördersätze für Wärmepumpen reichen bis zu 40 %, während kleine Einzelmaßnahmen wie der Austausch von Fenstern nur 15 % der Kosten decken (20 %, wenn sie mit einem Energieberater durchgeführt werden).

nahmen konfrontiert ist, was derzeit als Engpass auf dem Weg zu einer schnelleren Dekarbonisierung des Sektors angesehen wird. Angesichts dieses Fachkräftemangels könnte die Betonung der Schaffung von Arbeitsplätzen als positives Kriterium ohne unterstützende Maßnahmen zur Förderung der Aus- und Weiterbildung von Fachkräften dazu führen, dass ex-ante unrealistische Umsetzungsstrategien entwickelt werden.

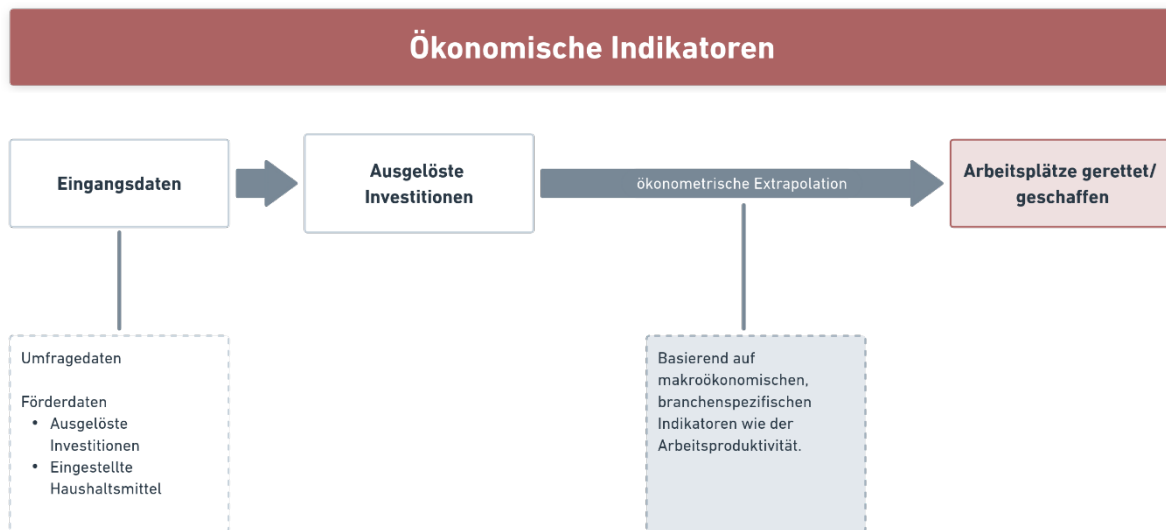


Abbildung 5: Berechnung der verursachten privaten Investitionen und der geschaffenen Arbeitsplätze.
Quelle: Eigene Darstellung

5.2.3. Verteilungseffekte

In den veröffentlichten Methodikleitfäden wird den sozioökonomischen Auswirkungen nur sehr wenig Beachtung geschenkt. Die Verteilungswirkungen politischer Optionen werden in den derzeit veröffentlichten Evaluierungen und Bewertungen meist nicht berücksichtigt [Interview 6, 8, 11, 12]. Die jüngsten BEG-Evaluierungen haben „soziale Aspekte“ einbezogen, die einige Informationen über die Empfänger von Fördermitteln, einschließlich Nettomonatsgehalt, Alter und Bildungsniveau, umfassen. Wie diese Informationen erhoben und ausgewertet werden, ist derzeit unklar. Die jüngste Ergänzung ist ermutigend, da die Ziel- und Einkommensgruppen, die von der steuerlichen Unterstützung profitiert haben, eine große Lücke in der Evaluierung der Bundesausgabenprogramme darstellen. In der Vergangenheit wurden diese Informationen von den Bundesbehörden, die die Daten erheben, nicht erfasst, so dass diese Evaluierungsdimension ausgeklammert wurde [Interview 6]. Die jüngste BEG-Evaluierung zeigt, dass die größten Nutznießenden die höchsten Einkommensgruppen sind. In der untersten Einkommensgruppe

(bis 1.000 Euro Nettomonatsgehalt) wurde die Förderung nicht in Anspruch genommen, in der Gruppe 1.000-2.000 Euro nur zu 8 %.

Die indirekten Auswirkungen dieser methodischen Verzerrung begünstigen vermutlich regressiv Ergebnisse. Es gibt keine direkten Subventionen für sozioökonomische Gruppen, die einen großen Teil ihres Einkommens für Heizung ausgeben. Dies hat erhebliche Auswirkungen auf die Gestaltung der Maßnahmen. Aufgrund der derzeit schlechten Wärmedämmung ist nicht der gesamte Wohnungsbestand für die Installation von Wärmepumpen ohne weitere Investitionen in die Energieeffizienz geeignet. Ohne Förderung bedeutet dies, dass diejenigen, die es sich nicht leisten können, die Sanierungsanforderungen zu erfüllen, von der Wärmepumpenförderung ausgeschlossen werden könnten. Dies könnte dazu führen, dass die Wärmepumpenförderung ungewollt regressiv ausgerichtet ist [Interview 6], obwohl sie derzeit als neutraler Mechanismus konzipiert ist. Zweitens wirkt sich die Kohlenstoffbepreisung viel stärker auf die Gebäude mit der schlechtesten Leistung aus und kann sich daher regressiver auswirken als bei wohlhabenden Haushalten, die kostspielige Renovierungen durchgeführt haben. Drittens: Ohne eine gezielte Unterstützung einkommensschwacher Gruppen bei der Steigerung der Effizienz ihrer Häuser werden die steigenden Regulierungsstandards im GEG die Kosten denjenigen aufbürden, die am wenigsten in der Lage sind zu zahlen, was zu weiteren regressiven Ergebnissen führt.

Dies spiegelt einen allgemeineren Ansatz in der deutschen Bürokratie und Politik wider, bei dem Klimapolitik und Sozialpolitik getrennt behandelt wurden. Die Klimapolitik hat sich in erster Linie auf die Verringerung des Energiebedarfs und die Reduzierung von Treibhausgasen konzentriert sowie darauf, diese Ziele kosteneffizient zu erreichen. Die Sozialpolitik wurde als separates Thema betrachtet, und erst in jüngster Zeit gab es mehr Bemühungen, diesen Politikbereich zu integrieren [Interview 8].

Auf EU-Ebene wurden vor kurzem die Anforderungen erhöht, sich sektorübergreifend auf die sozialen Aspekte der Energiepolitik zu konzentrieren. Dieser Schwerpunkt der EU hat zu einer expliziteren Berücksichtigung der sozioökonomischen Auswirkungen (Energiearmut usw.) geführt, und dieser Indikator muss in die nächste Berichtsrunde aufgenommen werden. Andere Länder, darunter Frankreich und das Vereinigte Königreich, haben der Energiearmut und den sozialen Auswirkungen der Maßnahmengestaltung

mehr Aufmerksamkeit gewidmet, einschließlich der gezielten Ausrichtung von Programmen auf besonders gefährdete Gruppen. Dies ist wahrscheinlich auf die geringere Qualität des Gebäudebestands zurückzuführen, die dazu geführt haben, dass die sozialen Auswirkungen der Energiepolitik über einen längeren Zeitraum hinweg stärker ins Blickfeld gerückt sind. Deutschland hatte sich, wie Schweden, nicht auf diesen Bereich konzentriert, aber mit den steigenden Energiekosten und der Umsetzung des BEHG sind diese Überlegungen in letzter Zeit sehr viel deutlicher geworden [Interview 8]. Darüber hinaus werden die Einführung des BEHG und die geplante Umsetzung von Mindesteffizienzstandards gesamtwirtschaftliche Auswirkungen mit Verteilungseffekten für sozioökonomische Gruppen haben. Dementsprechend ist eine Aktualisierung der Evaluierungsverfahren zur umfassenden Berücksichtigung dieser Dimensionen unerlässlich.

Wir schlagen vor, die sozioökonomischen Auswirkungen in der nächsten Runde der Evaluierungen stärker zu berücksichtigen und die Unterstützung ausdrücklich auf die am stärksten Betroffenen auszurichten. Die EU hat kürzlich die Aufnahme von Berichtspflichten für sozioökonomische Aspekte in die Evaluierungen vorgeschrieben [Interview 8]. Die jüngsten BEG-Evaluierungen enthalten einige zielgerichtete Aspekte, wie z.B. das Nettoeinkommen der Begünstigten, aber dies sollte erweitert werden, um die Senkung der Energiekosten im Verhältnis zu den Einkommensgruppen zu schätzen. Jüngste Berichte deuten darauf hin, dass die derzeitige Ausrichtung der BEG-Subventionen im Allgemeinen wohlhabendere Gruppen begünstigt. Diese Programme sollten überdacht werden, um die Progressivität zu erhöhen. Bisher wurden die bestehenden Methodikleitfäden noch nicht aktualisiert und um Verteilungseffekte erweitert. Folglich gibt es bisher nur wenige Anhaltspunkte dafür, wie die Verteilung berechnet werden soll. Die Erweiterung der bestehenden Evaluierungsmethodik scheint aufgrund der bereits erwähnten Beschränkungen bei den Evaluierungsverfahren und der Datenverfügbarkeit eine große Herausforderung darzustellen (siehe auch Abschnitt 5.3).

5.2.4. Akzeptanz

Akzeptanz wird nur im engeren Sinne als die Zufriedenheit der Fördermittelempfänger mit den installierten Maßnahmen betrachtet. Mittels Befragung werden die Fördermittelempfänger nach der Zufriedenheit mit den installierten Maßnahmen befragt. Dies liefert zwar einige Informationen zu einer Dimension der Akzeptanz, umfasst aber nur eine

Stichprobe von Bewilligungsempfängern. Über die Akzeptanz in der Allgemeinbevölkerung lassen sich damit keine verwertbaren Aussagen treffen.

Die Akzeptanz der Bevölkerung ist bei Ex-post-Evaluierungen keine explizite Dimension. Die Akzeptanz wird in den Leitfäden in Form von weichen Empfehlungen berücksichtigt, aber nicht explizit in evaluativen Dimensionen. Beispielsweise wird in den Leitfäden empfohlen, Handwerker als Multiplikator von Maßnahmen einzusetzen, da die Bürger Handwerkern mehr Vertrauen entgegenbringen als der Regierung.

Die Akzeptanz ist von der Ex-ante-Methodik ausgeschlossen. Eine gewisse Berücksichtigung der Akzeptanz in der Ex-ante-Evaluierungsmethodik ist zwar schwierig zu bewerten, könnte aber aufgenommen werden. Dies könnte dazu beitragen, gezielte Maßnahmen zu ermitteln, die zur Förderung von Instrumente erforderlich sein könnten, z.B. gezielte Informationskampagnen.

Die Akzeptanz ist nicht einfach zu bewerten, da es keine kontrafaktische Gruppe gibt. Die Ex-post-Evaluierung einer Stichprobe von Multiplikatoren repräsentiert nicht die breitere Population derjenigen, die die Maßnahmen wahrscheinlich nicht annehmen oder umsetzen werden. Selbst in der breiteren akademischen Forschung wird die Akzeptanz üblicherweise auf der Grundlage der Antworten auf nationale Umfragen und Meinungsumfragen geschätzt (Levi, 2021). Diese Art von Forschung basiert häufig auf einem einzigen politischen Instrumententyp, berücksichtigt aber nicht die Stringenz der Instrumente, die Wechselwirkungen mit anderen Instrumenten oder praktische Fragen wie die Unannehmlichkeiten und Störungen bei der Installation von Maßnahmen. Weitere Forschung in diesem Bereich ist eine kritische Lücke im derzeitigen Wissensstand, und ein besseres Verständnis ist notwendig, um eine schnellere Annahme und mehr Fortschritte in der Klimapolitik zu ermöglichen.

Die Akzeptanz in diesem Sektor hängt nicht immer von den politischen Instrumenten oder Programmen ab, sondern auch von den ergänzenden Maßnahmen, die erforderlich sind, um eine wirksamere Verwaltung zu ermöglichen. Wie bereits erörtert, sind der Datenschutz und der Einsatz intelligenter Messgeräte sowie die Echtzeitmessung des Energieverbrauchs wichtige Reformen, die die Zuverlässigkeit der Evaluierungen von Maßnahmen in diesem Sektor erheblich steigern könnten. Akzeptanzprobleme im Zusammenhang mit dem mangelnden Vertrauen in die Regierung und die Verwendung von

Daten waren jedoch häufig ein Hindernis für eine breitere Einführung von intelligenten Messgeräten in Haushalten (Bugden und Stedman, 2019).

5.2.5. Governance

Die Methodikleitfäden für Evaluierungen umfassen derzeit keine umfassenderen Aspekte der Governance-Anforderungen, wie Kapazitäten, Durchsetzung und Einhaltungssquoten. Folglich sind diese Aspekte vom Anwendungsbereich der Evaluierungen ausgeschlossen. In Anbetracht der potenziellen Bedeutung von Governance-Defiziten für die Wirksamkeit von Instrumenten und Programmen wäre es jedoch von Vorteil, sie ausdrücklich in den Evaluierungsprozess einzubeziehen. Unabhängige Berichte, die speziell diese Fragen untersuchen, z.B. im Hinblick auf die Wirksamkeit der Durchsetzung, können wertvolle Erkenntnisse liefern. Daher würde die Einbeziehung von Governance-Defiziten und ihren potenziellen Auswirkungen in den Evaluierungsrahmen das Verständnis für die allgemeine Wirksamkeit und Effizienz von Instrumenten und Maßnahmen verbessern. Es wird empfohlen, diese Praxis zu übernehmen, um eine umfassendere Evaluation von Programmen zu gewährleisten.

Die Governance-Anforderungen werden zum Teil über die Verwaltungskosten, die den Bundesbehörden für BEG-Maßnahmen entstehen, in den Ex-post-Evaluierungen berücksichtigt. Der Leitfaden sieht vor, dass die Verwaltungskosten der Bundesbehörden, die mit der Durchführung des BEG beauftragt sind, erfasst und bewertet werden sollen. Die Beschreibung der Verwaltungskosten umfasst insbesondere nicht die Verwaltung oder den Vollzug von Regulierungsmaßnahmen für die Bundesländer und die Landkreise.

In unserer Stichprobe von veröffentlichten Bewertungen wurden die Verwaltungskosten in den Berichten nicht berechnet. Obwohl die Leitfäden für Evaluierungen die Berücksichtigung von Verwaltungskosten vorsehen, werden diese Kosten in der Praxis häufig aus Ex-post-Evaluierungen ausgeklammert, da es schwierig ist, sie genau zu berechnen. Obwohl die Verwaltungskosten bei der Gesamtbewertung der Kosteneffizienz von Energieeffizienzprogrammen eine wichtige Rolle spielen, kann ihr Ausschluss aus den Ex-post-Evaluierungen das umfassende Verständnis der Effizienz des Programms einschrän-

ken. Es ist wichtig, diese Einschränkung zu erkennen und Möglichkeiten zur Verbesserung der Methodik für die Erfassung und Einbeziehung von Verwaltungskosten in den Evaluierungsprozess zu untersuchen, um eine genauere Bewertung zu gewährleisten.

Die Empfänger von Subventionen sind verpflichtet, ein Feedback über ihre Zufriedenheit mit dem Verwaltungsprozess abzugeben. Dies gibt Aufschluss über die Effizienz der föderalen Agenturen bei ihren Verwaltungsprozessen und der Installateure. Die gewonnenen Informationen sind jedoch aus der Sicht des Empfängers subjektiv. Die Daten können einer positiven Verzerrung unterliegen, da die Antworten derjenigen, die den Prozess abgeschlossen haben, zustimmender sind als die derjenigen, die ihn nicht abgeschlossen haben. Als kontrafaktische Größe werden auch diejenigen befragt, die eine Storno-Befragung beantragt, aber nicht abgeschlossen haben (KfW-433, S. 5-7), aber in einigen Fällen ist die Rücklaufquote gering. Ist die Anzahl der Fälle zu niedrig, ist eine Hochrechnung nicht immer möglich und wird daher von den Auswertungen ausgeschlossen.

5.3. Daten

Die Genauigkeit der Evaluierung hängt von der Zuverlässigkeit der Eingabedaten ab.

Die Datenverfügbarkeit hat großen Einfluss auf die Zuverlässigkeit der Evaluierung. Zuverlässige Eingabedaten in Bezug auf den Gebäudebestand, die derzeit verwendeten Energieträger und die Effizienz der durchgeführten Maßnahmen sind von entscheidender Bedeutung, um sicherzustellen, dass Energieeinsparungen und THG-Einsparung korrekt berechnet werden (Abschnitt 5.4). In diesem Abschnitt werden die wichtigsten Fragen im Zusammenhang mit Daten behandelt: Qualität, Berichterstattung, Schutz, Durchsetzung und Einrichtung einer Datenbank.

5.3.1. Datenqualität

Die Subjektivität der Anwendung von Datenqualitätsstandards im Evaluierungsprozess kann zu Unstimmigkeiten und Ungenauigkeiten in den Ergebnissen führen. In den Leitfäden werden zwar verschiedene Kategorien der Datenqualität umrissen, doch fehlt es an spezifischen Leitfäden, wie die Datenqualität für eine konkrete Evaluierung zu beurteilen ist. Diese Unklarheit lässt Raum für subjektive Urteile und Entscheidungen über

die Qualität der Daten, was letztlich zu uneinheitlichen und ungenauen Ergebnissen führen kann. Der Evaluierungsprozess kombiniert die Verwendung von technischen Daten und Umfrageantworten, aber bei mehreren wichtigen Kriterien fehlen technische Daten, was die Vollständigkeit und Zuverlässigkeit der Bewertung beeinträchtigen kann.



Abbildung 6: Visualisierung der Datengüte/Belastbarkeit. Quelle: Eigene Darstellung, adaptiert von Schlomann et al. (2020).

Eine weitere potenzielle Quelle für Verzerrungen ergibt sich aus der an der Bewertung beteiligten Stichprobengruppe. Die Selbstauskunft birgt die Möglichkeit subjektiver Interpretationen und Verzerrungen in den erhobenen Daten. Darüber hinaus kann eine niedrige Rücklaufquote von nicht ausgefüllten Fragebögen die Repräsentativität und Zuverlässigkeit der Daten weiter beeinträchtigen.

Der Datenzugang und Fragen der Qualität der Aufzeichnungen werden in den Evaluierungsleitfäden nicht ausdrücklich erörtert. Während die Qualität der Daten im Hinblick auf ihre Verfügbarkeit bewertet wird, wird nicht ausdrücklich auf Fragen der Zugänglichkeit und der Qualität der verfügbaren Daten eingegangen. Im Folgenden werden einige dieser Fragen explizit erörtert.

5.3.2. Qualität der Datenberichterstattung und Zugang zu den Daten in den Bundesagenturen

Die Daten über den Energieverbrauch des Empfängers werden von den Bundesbehörden für die Beantragung von Fördermechanismen erhoben, jedoch nicht anonymisiert.

Aufgrund der fehlenden Anonymisierung fallen diese Daten in den Geltungsbereich des Datenschutzes und sind nicht öffentlich zugänglich. Infolgedessen sind die Möglichkeiten zur Durchführung von Genauigkeits- und Gültigkeitsprüfungen begrenzt und die administrative Rechenschaftspflicht eingeschränkt. Da es kaum Gültigkeitsprüfungen oder Audits gibt, kann die Qualität der erfassten Daten beeinträchtigt werden. Beispielsweise können die erfassten Daten von geringer Qualität sein, und es gibt Fälle, in denen Gebäude mit unplausiblen Größen von nur 1 m² erfasst werden [Interview 7]. Diese Bedenken wurden von Beamten [Interview 3, 4] und Beratern [Interview 6, 7, 10, 12] geäußert. Dies stellt eine große Herausforderung bei der Durchführung von Evaluierungen dar, da die Beratungsunternehmen Daten von geringer Qualität auf Fehler überprüfen müssen. Wie diese Fehler korrigiert werden, ist möglicherweise nicht transparent oder standardisiert, was die Replizierbarkeit einschränken kann (Abschnitt 5.4).

Die Qualität der Daten zu den finanziellen Aspekten ist in der Regel hoch, und die Menge der Daten zu den durchgeführten Maßnahmen ist überwältigend. Wenn man sich die Maßnahmen im Detail anschaut, gibt es viele Details, die nicht verfügbar sind. So weiß man meist nicht, wie der Zustand vor der Sanierung war und wie er danach sein wird. Es gibt Dinge, die wir aus den Daten nicht herauslesen können. Manchmal befinden sich solche Daten in einer Datei, aber sie werden nicht als elektronische Daten gesammelt und gespeichert. [Interview 7].

Die Daten werden den beauftragten Beratern in anonymisierter Form zur Verfügung gestellt. Nach Erhalt der Daten führen die Gutachter eine erste Datenprüfung auf fehlende/unplausible Daten durch. Diese werden mit verschiedenen Methoden korrigiert. Was getan wird und wie es getan wird, wird zumindest für den Arbeitgeber transparent gemacht - ein Verfahren, das bei Evaluatoren üblich ist [Interview 7]. Allerdings ist dies nicht transparent oder öffentlich zugänglich, was die Replikation einschränkt. Darüber hinaus werden häufig weitere Erhebungen bei den Geförderten durchgeführt, z.B. um Basiswerte zu erhalten oder um Gebäudeparks detaillierter modellieren zu können [Interview 7].

Die fehlende Digitalisierung der Aufzeichnungen ist ein weiteres Hindernis für die Beschleunigung der Evaluierungsprozesse. Die gesammelten Daten sind nicht digitalisiert

und werden in physischen Akten aufbewahrt, so dass die Berater Zugang zu den Informationen beantragen müssen, um sie einzusehen. Dieser Prozess kann zu Verzögerungen führen und die zeitnahe Analyse und Bewertung der Daten behindern: „Finanzierungsdaten, die man erst nach vier Monaten und 25 E-Mails erhält“ [Interview 9]. Dies war ein von den Beratern häufig genanntes Problem [Interview 7, 10, 12].

Die Erfassung von Kennzahlen und Kostenstrukturen variiert zwischen Behörden. Im Vergleich zum BAFA sind die Kostenstrukturen der KfW generell höher. Die KfW verwendet Berechnungsmodelle, die besser für umfassende Sanierungsprojekte geeignet sind, die im Vergleich zu den vom BAFA abgedeckten Einzelmaßnahmen tendenziell komplexer sind. Die Programme des BAFA sind auf einen breiteren Markt ausgerichtet und vergleichsweise einfach zu handhaben. Diese Unterschiede in der Herangehensweise und Komplexität sind kein Hinweis auf Präferenzen, sondern spiegeln vielmehr den spezifischen Fokus und Umfang der Förderaktivitäten der jeweiligen Institution wider [Interview 4].

Die Ermöglichung der Überprüfung und Validierung der aufgezeichneten Daten kann sich auf die Zuverlässigkeit und Glaubwürdigkeit der Evaluierungsergebnisse auswirken. Um diese Probleme zu lösen, müssen Lösungen gefunden werden, die ein Gleichgewicht zwischen den Anforderungen des Datenschutzes und dem Bedürfnis nach Transparenz, Genauigkeit und Gültigkeit der Evaluierungsprozesse herstellen. Dazu könnte die Untersuchung von Anonymisierungsmethoden gehören, um Daten für Test- und Analysezwecke zugänglich zu machen, ohne die Privatsphäre zu gefährden. Darüber hinaus könnten die Digitalisierung der Daten und die Einführung benutzerfreundlicher Plattformen für den Datenzugang den Evaluierungsprozess rationalisieren und die Effizienz verbessern.

5.3.3. Datenschutz und Durchsetzung von Rechtsinstrumenten

Die Evaluierung von Vorschriften im Hinblick auf das Erreichen von Energieeffizienzzielen wird durch mehrere zentrale Herausforderungen im Zusammenhang mit der Datenverfügbarkeit, der Durchsetzung und der Rechenschaftspflicht behindert. Erstens erschwert das Fehlen von Daten über den Energieverbrauch vor der Umsetzung von

Energieeffizienznormen, wie im vorigen Absatz erwähnt, die Bewertung des Ausgangsenergieverbrauchs und die Ermittlung der Auswirkungen der Vorschriften. Darüber hinaus mangelt es an Daten über die Auswirkungen der Verordnungen nach ihrer Umsetzung, was vor allem darauf zurückzuführen ist, dass es keine Berichterstattungspflicht gibt. Ohne umfassende Daten über Energieverbrauchsmuster und Leistungsindikatoren ist es schwierig, die Wirksamkeit von Regulierungsmaßnahmen genau einzuschätzen.

In ihrer Kernfrage geht es um die Bewertung, und eine Bewertung setzt Daten voraus, wenn man sie nicht nur in einer Blase durchführen will. Ich denke, das ist einfach der Kern der Sache. Wir haben keine-, wir investieren, um es brutal auszudrücken, aber ich glaube, ich habe schon in der letzten Runde mit Ihnen gesagt: Wir investieren Milliarden und kennen den Ist-Zustand nicht wirklich und damit natürlich auch nicht die Wirkung. Also selbst wenn ich im Nachhinein feststellen würde, dass der Verbrauch jetzt so und so wäre. Ich weiß ja nicht einmal, woher ich komme. Ich muss eine tatsächliche [Kontrolle] haben und dann ein Ziel und dann eine Erfüllung. [Interview 4].

Eines der wichtigsten Probleme, das die Wirksamkeit von regulatorischen Instrumenten untergräbt, ist der mangelhafte Vollzug. Die Nichteinhaltung von Vorschriften im Bausektor ist allgemein bekannt. Schätzungen gehen von einer Nichteinhaltungsquote von bis zu 25 % oder möglicherweise sogar noch höher aus (Lu et al., 2022). Dies macht es äußerst schwierig, die tatsächlichen Auswirkungen von regulatorischen Instrumenten auf die Energieeffizienz zu beurteilen. Selbst wenn Reformen zur Verbesserung der Datenbereitstellung und des Datenzugangs durchgeführt werden, hängt die Wirksamkeit dieser Maßnahmen in hohem Maße von einer robusten Aufsichtsbehörde ab. Die derzeitigen Einschränkungen bei der Bereitstellung und Weitergabe von Daten, einschließlich der Schwierigkeiten der Bundesregierung beim Zugriff auf regionale Daten, tragen jedoch zu einem Mangel an Rechenschaftspflicht bei. Dies untergräbt folglich die Glaubwürdigkeit und Wirksamkeit des für die Durchsetzung der Vorschriften zuständigen Inspektionssystems.

In Deutschland gibt es keine Bauaufsicht. Wenn ich also ein Gebäude habe und zehn Prozent oder mehr eines Bauteils, zum Beispiel der Außenwand, reparieren will, muss ich es energetisch auf Neubaustandard bringen, so ist es gesetzlich vorgeschrieben. Aber niemand kontrolliert das, niemand kümmert sich darum. [Interview 14].

Es werden mehr Mittel für eine robuste Aufsicht auf Länderebene und eine bessere Koordinierung mit der Bundesregierung benötigt. Um die Durchsetzung der Vorschriften zu verbessern, müssen die Mittel, die Kompetenzen und die Ausbildung auf regionaler Ebene (Länder) aufgestockt und die vertikale Koordinierung zwischen den verschiedenen Regierungsebenen verbessert werden. Die Bewältigung dieser Aufgaben ist jedoch kein einfaches Unterfangen, was die bestehenden Mängel erklären könnte. Nichtsdestotrotz können das Erkennen und Angehen dieser administrativen Probleme, die Reform der Datenbereitstellung und des Datenzugangs, die Stärkung der Rechenschaftspflicht und die Durchführung von Evaluierungen zur Bewertung der Leistung von Vorschriften einen erheblichen Einfluss auf die Fortschritte bei der Dekarbonisierung des Bausektors haben. Durch die Verbesserung der Datenverfügbarkeit, der Durchsetzungsmechanismen und der Strukturen für die Rechenschaftspflicht können die politischen Entscheidungsträger wirksamere und wirkungsvollere regulatorische Maßnahmen fördern, die zu den gewünschten Ergebnissen beitragen.

Die Überprüfung und Kontrolle der Regulierungspolitik ist sehr schwierig. Das müssen die Länder machen und nicht der Bund. Es gibt sehr wenig, erstaunlich wenig Daten. Wir machen gerade das Projekt zur Entwicklung der nächsten Generation des Abrechnungs-codes, letztes und dieses Jahr, das ist ein riesiges Projekt, und es gibt so wenig Daten, auf denen man aufbauen kann. Wir können die allgemeinen Trends erkennen, aber wir können weder regionalisierte Daten noch [Einhaltungsquoten] sehen. Wie kann man also die Regulierung bewerten? Das ist noch viel schwieriger. [Interview 9]

5.3.4. Einrichtung einer Gebäudedatenbank

Die Verfügbarkeit von Daten stellt ein erhebliches Hindernis für eine qualitativ bessere Evaluierung dar. Die begrenzte Verfügbarkeit von Daten über den aktuellen Gebäudebestand stellt ein erhebliches Hindernis für eine genaue Bewertung der durch Energieeffizienzmaßnahmen eingesparten Energie dar. Dieses Defizit wurde kürzlich in einem Bericht des Expertenrats für Klimafragen (ERK) festgestellt:

Die Verfügbarkeit von öffentlich zugänglichen Daten im Gebäudesektor ist begrenzt. In Deutschland gibt es zwar amtliche Statistiken sowie wissenschaftliche und wirtschaftliche Statistiken wie Verbandszahlen oder Förderstatistiken, die Auskunft über einzelne Merkmale des Gebäudebestands geben. Veränderungen im Gebäudebestand im Hinblick auf die

baulichen Eigenschaften der Heizungsanlage und die energetische Qualität der Gebäudehülle werden jedoch nicht regelmäßig amtlich erhoben. Daher gibt es Datenlücken, die eine kontinuierliche Beobachtung des Strukturwandels erschweren. So können wichtige Indikatoren oder Daten zur energetischen Sanierungsrate und zum energetischen Zustand des Gebäudebestands oder zu Energieeffizienzlabeln nicht ausreichend erfasst werden. (ERK 2022; S.24)

Deutschland verfügt im Gegensatz zu den meisten EU-Ländern nicht über eine Gebäudedatenbank. Mit Ausnahme von Deutschland und Rumänien unterhalten alle anderen Länder eine Energieausweisdatenbank und nutzen die Daten daraus, um Evaluierungen und Aussagen über ihren Gebäudebestand zu treffen (*BfEE-data-quality-building-sector*, S. 6). Diese Datenbanken dienen als wertvolle Ressourcen für die Sammlung von Informationen über die Energieeffizienz von Gebäuden und ermöglichen es politischen Entscheidungstragenden, Forschenden und Interessengruppen, die Energieeffizienz der bebauten Umwelt zu analysieren und zu bewerten. Das derzeitige Fehlen einer zentralen Energieausweisdatenbank in Deutschland schränkt die Möglichkeit ein, umfassende Aussagen über die Energieeffizienz des Gebäudebestands zu treffen.

Um eine bessere Evaluierung und Analyse zu ermöglichen, ist die Einrichtung einer umfassenden Datenbank von entscheidender Bedeutung. Die Einrichtung einer Datenbank ist eine Priorität, um bessere Evaluierungsverfahren zu ermöglichen [Interview 4, 6, 7, 10, 11, 12, 14]. Diese Datenbank sollte Informationen über den Gebäudebestand enthalten, einschließlich der Indikatoren für die Gesamtenergieeffizienz, Effizienzklassen und anderer relevanter Merkmale. Wenn diese Datenbank öffentlich zugänglich gemacht wird, sorgt sie für Transparenz und erleichtert die Forschung, die Entwicklung von Strategien und die Überwachung von Energieeffizienzmaßnahmen. In Übereinstimmung mit unseren Ergebnissen hat der Expertenrat für Klimafragen (ERK) in seinem Bericht ebenfalls eine Gebäudedatenbank empfohlen (ERK 2022; S.24)

Die verschiedentlich diskutierte Einrichtung eines bundesweiten Gebäude- und Wohnungsregisters würde wesentlich dazu beitragen, eine vollständige Transparenz über den Sanierungsstand und die technische Ausstattung des Gebäudesektors zu erhalten und damit einen wesentlichen Beitrag zu einem effektiven Monitoring der für die Emissionsentwicklung relevanten Ursachen leisten. So gibt es in Österreich bereits seit 2004 ein Gebäude-

und Wohnungsregister. Darüber hinaus gibt es in Österreich eine Energieausweisdatenbank. Mit Hilfe dieser Daten können neben wohnungspolitischen und raumplanerischen Fragen auch gezieltere Maßnahmen zur Erreichung der Klimaziele entwickelt werden. Die Erkenntnisse aus einer solchen Datenbank können z.B. direkte Auswirkungen auf Förderbestände und -höhen haben (Statistik Austria 2022).

Die Einführung und Durchsetzung von Mindesteffizienzstandards würde die Verfügbarkeit und den Zugang zu Daten über den Gebäudebestand erheblich verbessern.

Deutschland wird im Rahmen der neuen Reform der EU-Kommission, die in den Fit-for-55-Vorschlägen enthalten ist, verpflichtet sein, Mindesteffizienzstandards einzuführen. Die Einführung von Mindesteffizienzstandards wird erhebliche Änderungen bei der Durchsetzung und der gemeinsamen Nutzung von Daten für Vorschriften in Deutschland erfordern. Wirksame Durchsetzungsmechanismen sind erforderlich, um die Einhaltung von Energieeffizienzstandards zu gewährleisten und die Umgehung von Vorschriften zu verhindern. Darüber hinaus ist ein verbesserter Datenaustausch zwischen den Ländern und der Bundesregierung von entscheidender Bedeutung für ein umfassendes Verständnis der Gesamtenergieeffizienz von Gebäuden im ganzen Land. Im Laufe der Zeit, wenn die Mindesteffizienzstandards stärker vereinheitlicht werden, können diese als Maßstab für die Bewertung von Energieeinsparungen dienen.

Mit der EPBD, der Europäischen Gebäuderichtlinie, müssen wir eine Gebäudedatenbank einführen. Das ist sehr hilfreich, denn das wäre ein automatischer und statistisch gesehen ein wunderbarer Ausgangspunkt für Bewertungen, denn dann kann man wirklich in Echtzeit sehen, was passiert. Es wird noch einige Jahre dauern, bis wir das wirklich haben. [Interview 9]

Es könnte eine obligatorische Inspektion der bestehenden Gebäude und Leistungsstandards eingeführt und darüber berichtet werden. Eine mögliche Lösung zur Behebung dieser Datenlücke ist die Durchführung von Inspektionen des Wohnungsbestands. Dabei würde die Gesamtenergieeffizienz von Gebäuden systematisch bewertet, um relevante Daten zu sammeln. Ähnliche Inspektionen sind derzeit für Gasheizkessel vorgeschrieben, so dass die Ausbildung und die Fähigkeiten des bestehenden Inspektionssystems auf die Bewertung der gesamten Immobilie ausgeweitet werden könnten. Um die

Einhaltung der Datenschutzbestimmungen zu gewährleisten, könnten die gesammelten Daten anonymisiert werden, um Konflikte mit dem Datenschutz zu vermeiden.

Nach den geltenden deutschen Datenschutzgesetzen ist die Verwendung von Energieausweisen zum Aufbau einer Datenbank nicht zulässig. Diese Einschränkung schränkt die Durchführbarkeit von Methoden ein, die sich auf Energieausweisdatenbanken stützen. Aufgrund von Bedenken hinsichtlich des Schutzes personenbezogener Informationen können Daten auf individueller Ebene nicht abgerufen oder für Analysen verwendet werden. Infolgedessen können Daten nur in Form von Stichproben gespeichert und verwendet werden, was umfassende Evaluierungen und Analysen der Energieeffizienz im Gebäudesektor erschwert.

Alternativ könnte eine bessere Datenqualität durch Echtzeitüberwachung und eine bessere Zugänglichkeit der Daten von Energieversorgern ermöglicht werden. Eine genauere Überwachung der Gebäudestandards und des Energieverbrauchs in Echtzeit würde dazu beitragen, diese Lücke in den aktuellen Eingangsdaten zu schließen [Interview 3, 6, 7, 11, 12]. Die Beschleunigung der Installation von intelligenten Messgeräten in Wohngebäuden würde dazu beitragen, genauere Daten über den Energieverbrauch der Haushalte zu erhalten. Die öffentliche Akzeptanz ist nach wie vor ein entscheidender Faktor, aber eine aktuelle Umfrage hat ergeben, dass z.B. 41 % der Berliner Bürgerinnen und Bürger bereit sind, intelligente Zähler zu installieren⁴. Eine Reform der Datenschutzgesetze, so dass der Energieverbrauch anonym von den Energieversorgern zur Verfügung gestellt werden kann, könnte auch die Bereitstellung von Daten für Bewertungs- und Forschungszwecke verbessern.

Derzeit gibt es zwei Möglichkeiten, eine Gebäudedatenbank zu erstellen, ohne die Datenschutzgesetze zu reformieren. In Anbetracht der Beschränkungen der Energieausweise für den Aufbau von Datenbanken: (i) Analyse von Fernerkundungsdaten und (ii) Analyse von Daten aus mehrstufigen Stichproben. Die Fernerkundungsdatenanalyse nutzt die äußeren Merkmale von Gebäuden. Sie stützt sich auf verschiedene Datenquellen, darunter Satellitenbilder, Luftaufnahmen, Katasterkarten, Daten lokaler Behörden und offene Datenquellen. Dieser Ansatz bietet den Vorteil, dass mehrere Datenquellen

⁴ <https://www.bitkom.org/Presse/Presseinformation/Interesse-an-Smart-Metern-steigt>

genutzt werden können, um umfassende und detaillierte Informationen über Gebäude zu erhalten. Er enthält jedoch keine detaillierten Informationen über die innere Zusammensetzung von Gebäuden.

Das IWU (Institut Wohnen und Umwelt) hat mit Hilfe einer mehrstufigen Stichprobenanalyse eine deutsche Gebäudetypologie entwickelt. Im Rahmen früherer Projekte (ENOB:dataNWG) wurden umfangreiche Erhebungen des Wohngebäudebestands durchgeführt (BfEE-data-quality-building-sector, S. 39). Derzeit läuft ein neues Projekt, das sich auf die Erhebung des leichten Wohngebäudebestands konzentriert, bei dem die Renovierungsaktivitäten unabhängig von den Finanzierungsquellen erfasst werden. Dies liefert wertvolle Einblicke in die Dynamik von Gebäuderenovierungen und deren Zeitplan. Es ist jedoch anzumerken, dass diese Art von Forschung sowohl zeitaufwändig ist als auch leider nicht konsequent in die laufende Praxis integriert wird. Die im Rahmen dieser Projekte gesammelten Daten sind wertvoll, da sie ein tieferes Verständnis der zeitlichen Muster und Trends im Zusammenhang mit Gebäuderenovierungen ermöglichen. Durch die Erfassung von Informationen über Renovierungsaktivitäten unabhängig von den Finanzierungsquellen wird ein besserer Überblick über die breitere Landschaft ermöglicht und ein Licht auf die Motivationen und Triebkräfte hinter Renovierungsentscheidungen geworfen. Die Durchführung solcher Erhebungen und das Sammeln dieser Fülle von Informationen erfordert einen erheblichen Zeit- und Ressourcenaufwand.

Die derzeitige Laufzeit dieser Studien ist zu lang, um innerhalb eines kurzen Zeitrahmens genaue Daten zu liefern. Der Prozess der Datenerhebung, -analyse und -auswertung ist keine einfache Aufgabe, da sich das letzte Projekt über fünf Jahre erstreckte. Darüber hinaus besteht die Herausforderung darin, die Kontinuität der Studien im Laufe der Zeit zu gewährleisten. Es ist wichtig, diese Bemühungen zu verankern und zu institutionalisieren, um eine konsequente Überwachung der Gebäudesanierungsaktivitäten zu ermöglichen [Interview 6]. Dies würde ein umfassenderes Verständnis von Trends, Verschiebungen und den Auswirkungen politischer Maßnahmen über einen längeren Zeitraum ermöglichen. Dies ist bisher noch nicht geschehen, sondern erfolgt immer projektbezogen über einige Jahre, und dann gibt es eine Pause von ein paar Jahren. Bei ihnen beträgt der Abstand zwischen den Erhebungen fast zehn Jahre. Angesichts der Dringlichkeit des Klimaschutzes im Gebäudesektor ist das nicht häufig genug [Interview 6].

Die Vorschläge für das Gesetz zur Wärmeplanung und Dekarbonisierung von Wärmenetzen könnten zu einer Verbesserung der Datenqualität und des Zugangs führen.

Während der Neufassung des Gesetzes über die Wärmeplanung und die Dekarbonisierung von Wärmenetzen wurden Vorschläge zur Verbesserung der Datenqualität gemacht. Dazu gehörte die Erhebung von Daten über den Stromverbrauch über einen Zeitraum von drei Jahren sowie von Informationen über die Art der verwendeten Heizungssysteme. Diese Datenerhebungsmaßnahmen zielten darauf ab, wichtige Informationen über Energieverbrauchsmuster und die Art der eingesetzten Heiztechnologien zu sammeln. Aus dem Gesetzentwurf, der den Kommunen am 21. Juli 2023 vorgelegt wurde, geht jedoch hervor, dass die Datenanforderungen möglicherweise gelockert wurden (BDEW, 2023). Zum Zeitpunkt der Erstellung dieses Berichts ist unklar, wie das Gesetz in seiner endgültigen Fassung aussehen wird, aber die ursprünglichen Vorschläge zur Ausweitung der Datenerfassung würden sowohl die Qualität der Evaluierungsprozesse verbessern als auch die Einhaltung der EU-Anforderungen zur Einrichtung einer Datenbank im Rahmen der EPBD unterstützen.

5.4. Methoden

Nachdem wir den Umfang und die Daten der Evaluierungsprozesse erörtert haben, konzentrieren wir uns nun auf die Methoden. Wie bereits erwähnt, liegt das Hauptaugenmerk der Evaluierungsprozesse auf der Umwelteffektivität der Instrumente und Programme (Energieeinsparungen und THG-Einsparung) und ihrer Kosteneffektivität. Angesichts der Bedeutung dieser Kriterien gehen wir auf die Methoden zur Berechnung dieser beiden Indikatoren näher ein. Dabei achten wir insbesondere auf die Transparenz, Zuverlässigkeit und Replizierbarkeit dieser Methoden. Anschließend befassen wir uns mit den Interaktionseffekten zwischen den Instrumenten und mit der Frage, wie diese in den derzeitigen Evaluierungsverfahren berücksichtigt werden. Der Abschnitt schließt mit einigen Empfehlungen zur Verbesserung der Verfahren in diesen Bereichen.

5.4.1. End- und Primärenergieeinsparungen

Die endgültigen Energieeinsparungen werden als Differenz zwischen dem geschätzten Energiebedarf vor und nach der Durchführung der Maßnahme berechnet. Um die Auswirkungen von Energiesparmaßnahmen zu bewerten, ist der Ausgangspunkt eine Schätzung des Gesamtenergiebedarfs, bei der Faktoren wie Gebäudeeigenschaften, Belegung und energieverbrauchende Geräte berücksichtigt werden (Ex-post-Leitfaden, S. 56-57). Diese Schätzung liefert eine Ausgangsbasis, an der die Wirksamkeit der Maßnahmen gemessen werden kann. Anschließend wird bei der Bewertung der Energiebedarf nach der Umsetzung bestimmter Maßnahmen berechnet, wobei der Energieträgermix, Effizienzverbesserungen und Änderungen im Nutzerverhalten berücksichtigt werden. Durch den Vergleich der anfänglichen Schätzung des Energiebedarfs mit dem endgültigen Energiebedarf können die Auswirkungen der durchgeführten Maßnahmen auf die Energieeinsparungen und die Verringerung der Treibhausgasemissionen ermittelt werden.

Energetischer Ausgleich

Die derzeitige Methode zur Berechnung des Gesamtenergiebedarfs ist die energetische Gebäudebilanzierung und die Anwendung eines Gebäudemodells. Allerdings ist dieser Prozess derzeit intransparent, was sich in der Unklarheit der angewandten Methoden bei den codierten Ex-Post-Evaluierungen zeigt (KfW-EBS-WG-2018-2021, S. 9, 40; KfW-433, S. 97-98; Hocheffizienzpumpen, S. 15-16). Eine mögliche Sorge der Befragten ist, dass diese Anpassungen nicht standardisiert sind. Es gibt zwar eine vom IWU entwickelte vereinfachte Gleichung, aber es liegt im Ermessen des Bewerter, wie er diese von der Berechnung pro m² auf eine relativ große Gebäudefläche (d.h. 150 m²) anwendet [Interview 9, 12, 14]. Infolgedessen wird durch diesen Mangel an methodischer Klarheit der tatsächliche Energieverbrauch eines Gebäudetyps möglicherweise überschätzt, wobei die schlechtesten Gebäude eine Überschätzung von bis zu 40-50 % erfahren [Interview 12, 14], ein Ergebnis, das durch frühere Untersuchungen zum Prebound-Effekt bestätigt wird (Rosenow und Galvin, 2013; Sunikka-Blank und Galvin, 2012). Jüngste Arbeiten zur Gebäudesanierung haben sich mit diesem Thema auseinandergesetzt (Jagnow und Wolff, 2020) und könnten in die Überarbeitung der Evaluierungsmethodik einbezogen werden.

Die begrenzte Standardisierung erschwert die Vergleichbarkeit der Ergebnisse verschiedener Studien. Aufgrund der fehlenden Standardisierung müssen beide Messgrößen für die Vergleichbarkeit zwischen den Studien erfasst werden [Interview 9]. Um diese Ergebnisse zu interpretieren, müssen die von den verschiedenen Beratungsunternehmen angewandten Anpassungsfaktoren transparent gemacht werden. Dies mag zwar für technisch versierte Beratungsunternehmen möglich sein, verringert jedoch die Transparenz dieser Ergebnisse für politische Entscheidungsträger oder öffentlich zugängliche Veröffentlichungen.

Was wir jetzt machen, ist, dass wir ein Gebäudemodell anwenden, um die Einsparungen pro Gebäude zu berechnen. Und dann haben wir Normalisierungsfaktoren, die sogenannten „Bedarf-Verbrauch-Faktoren“, die für alle verschiedenen Gebäudekategorien abgeleitet wurden. Es gibt also Faktoren, die zu multiplizieren sind, um die berechneten Verbrauchseinsparungen in reale Verbrauchseinsparungen umzuwandeln. Wir kalibrieren das, und in der Regel geben wir beides an. Wenn man die Bewertungsdaten verwendet, um zu sehen, ob die Maßnahmen, die in Deutschland umgesetzt werden, um unsere Klimaziele zu erreichen, ausreichend sind, hilft es nicht, die berechneten Daten zu verwenden, sondern man muss sehen, was wirklich passiert. Wenn man sie hingegen mit anderen Studien vergleichen will, kann es hilfreich sein, auch die berechneten Daten zu haben, wenn man andere Normalisierungs- oder Korrekturfaktoren anwendet [Interview 9].

Endenergieeinsparungen

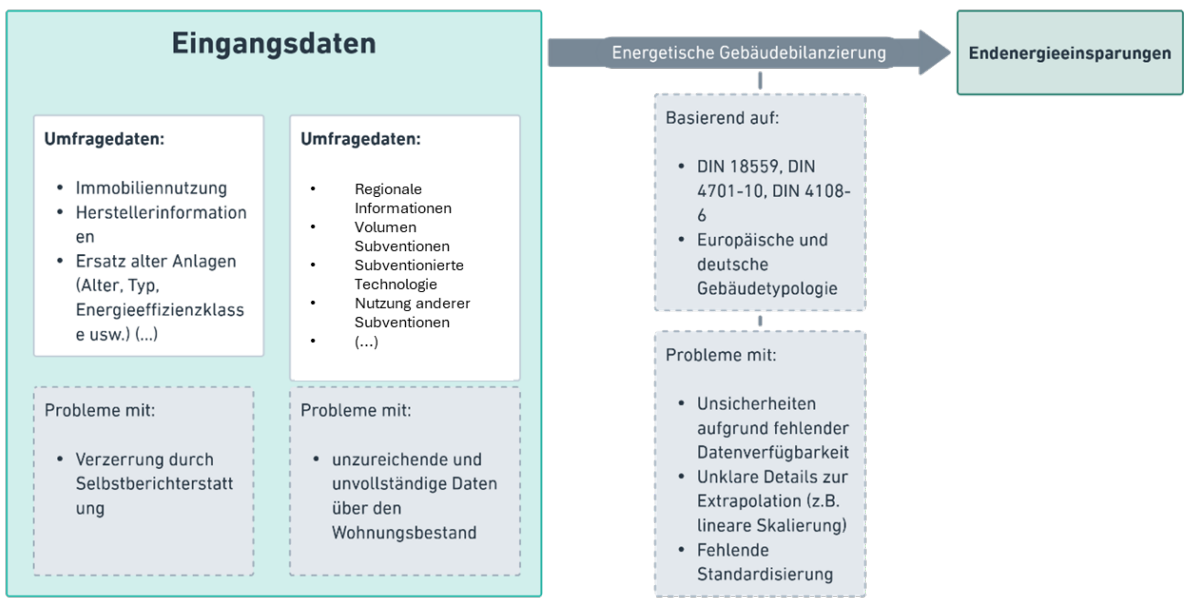


Abbildung 7: Berechnung der Endenergieeinsparungen. Quelle: Eigene Darstellung

Die vorhandenen Bewertungen geben keinen klaren Einblick in die Gewissheit oder Unsicherheit der endgültigen Energiebedarfsberechnungen. Ihnen fehlen Details wie Konfidenzintervalle oder Sensitivitätsanalysen. Stattdessen werden meist geschätzte Zahlen und Verweise auf DIN-Normen präsentiert, aber keine detaillierten Informationen über die Zuverlässigkeit oder Variabilität dieser Schätzungen gegeben (KfW-EBS-WG-2018-2021, S. 8-9). Diese Informationslücke erschwert ein umfassendes Verständnis der Genauigkeit und Robustheit der Energiebedarfsabschätzungen und erfordert daher mehr Transparenz bei ihrer Durchführung und Berichterstattung.

Die Anwendung dieser Evaluierungsmethoden ist bei den verschiedenen Beratungsunternehmen unterschiedlich ausgeprägt. Aufgrund von Problemen mit der Qualität der aufgezeichneten Daten sind die Beraterinnen und Berater verpflichtet, Berechnungs- und Anpassungsmethoden anzuwenden, aber die Art und Weise, wie diese angewendet werden, kann von Berater zu Berater variieren [Interview 9, 11, 12]. Diese Diskrepanz unterstreicht den Bedarf an größerer Transparenz und Standardisierung, insbesondere weil diese Berechnungen die Politik und die Ergebnisse der Maßnahmen erheblich beeinflussen, indem sie sich auf die Berechnung der endgültigen Energieeinsparungen auswirken.

*Im Allgemeinen gibt es einen Leitfaden für die Berechnung dieser Energie- und Kohlenstoff-
einsparungen. Und wir verwenden diesen Leitfaden, aber er passt nicht perfekt zu unse-
rem Projekt. [...] Wir haben die Daten für jeden einzelnen Fall. Wie immer sind die Daten
also nicht perfekt. Wir müssen uns also die Daten ansehen, und wenn es immer wieder feh-
lende Zahlen und offensichtlich falsche Zahlen gibt, versuchen wir, diese auszusortieren.
Und dann berechnen wir, und wir verwenden verschiedene Methoden für die Berechnung.
[...] Wir nehmen nicht nur die berechneten Werte, sondern wir versuchen, die tatsächliche
Effizienz zu schätzen. Aber die weicht meistens ein bisschen von der standardisierten Be-
rechnung ab [Interview 11].*

**Die Endenergieeinsparungen werden mit Hilfe des Primärenergiefaktors in Primär-
energieeinsparungen umgerechnet.** Der Primärenergiefaktor ist ein Koeffizient, der zur
Quantifizierung der Gesamtmenge an Primärenergiequellen (wie Kohle, Erdgas oder er-
neuerbare Energiequellen) verwendet wird, die zur Erzeugung einer Einheit End- oder
Nutzenergie erforderlich ist (Ex-post-Leitfaden, S. 18-19). Die Endenergieeinsparungen
werden durch Anwendung dieses Primärenergiefaktors in Primärenergieeinsparungen
umgerechnet. Die Berechnung basiert auf der Methodik des bestehenden Verfahrens zur
Ermittlung des Energiebedarfs. Daher wirken sich Ungenauigkeiten bei der Berechnung
der Energieeinsparungen – sei es bei der Anwendung der energetischen Bilanzierung
oder bei der Gültigkeit der Eingabedaten – direkt auf die resultierenden Primärenergie-
einsparungen und damit auf die THG-Einsparung aus.

5.4.2. Auswirkungen der Anpassung

**In den Leitfäden wird Effektbereinigung zur Korrektur von Brutto- auf Nettoeffekte be-
schrieben.** Dazu gehört die Berücksichtigung von Mitnahmeeffekten, Spill-over-Effekten
und Vorzieheffekten (Ex-post-Leitfaden, S. 86). Die methodische Anwendung der Effektb-
ereinigung ist in Abbildung 8 dargestellt.

**Mit den Auswirkungen auf die Gesamtkosten soll die Zusätzlichkeit der Maßnahme er-
fasst werden.** Mit dieser Anpassung soll beurteilt werden, ob die Maßnahmen auch ohne
den Zuschuss durchgeführt worden wären. Dies betrifft das, was Ökonomen als Zusätz-
lichkeit bezeichnen, und bezieht sich auf die Kosteneffizienz des steuerlichen Zuschusses.

Nicht genügend Daten über die allgemeine Bevölkerung für eine kontrafaktische

Gruppe. Es ist sehr wenig über Gruppen bekannt, die keine steuerliche Unterstützung in Anspruch nehmen oder die Maßnahmen ohne Unterstützung ergreifen [Interview 6, 5, 7, 12]. Dies macht es sehr schwierig, eine kontrafaktische Gruppe zu bilden, um den zusätzlichen Nutzen der steuerlichen Förderprogramme zu ermitteln.

Der ganze Bereich, was in Gebäuden passiert, die keine Fördermittel beantragen, usw., das ist einfach etwas, wo wir nichts wissen, wo eine permanente, kontinuierliche Überwachung wirklich sehr hilfreich wäre. [Interview 4].

Diese Effekte werden auf Basis von Befragungsdaten geschätzt. Alle einbezogenen Bereinigungseffekte basieren auf Befragungen von Teilnehmern der Maßnahme (KfW-EBS-WG-2018-2021, S. 60; KfW-433, S. 84). Für beliebige Effekte ist dies (zumindest für eine allgemeine Hochrechnung) problematisch, da die *Stichprobe* nicht die Gesamtbevölkerung widerspiegelt. Empfänger von Zuschüssen für teurere Sanierungen haben die Maßnahmen mit sehr viel höherer Wahrscheinlichkeit bereits in Betracht gezogen. Von den befragten Teilnehmern werden die Folgekosten auf bis zu 50 % geschätzt (KfW-EBS-WG-2018-2021, S. 63).

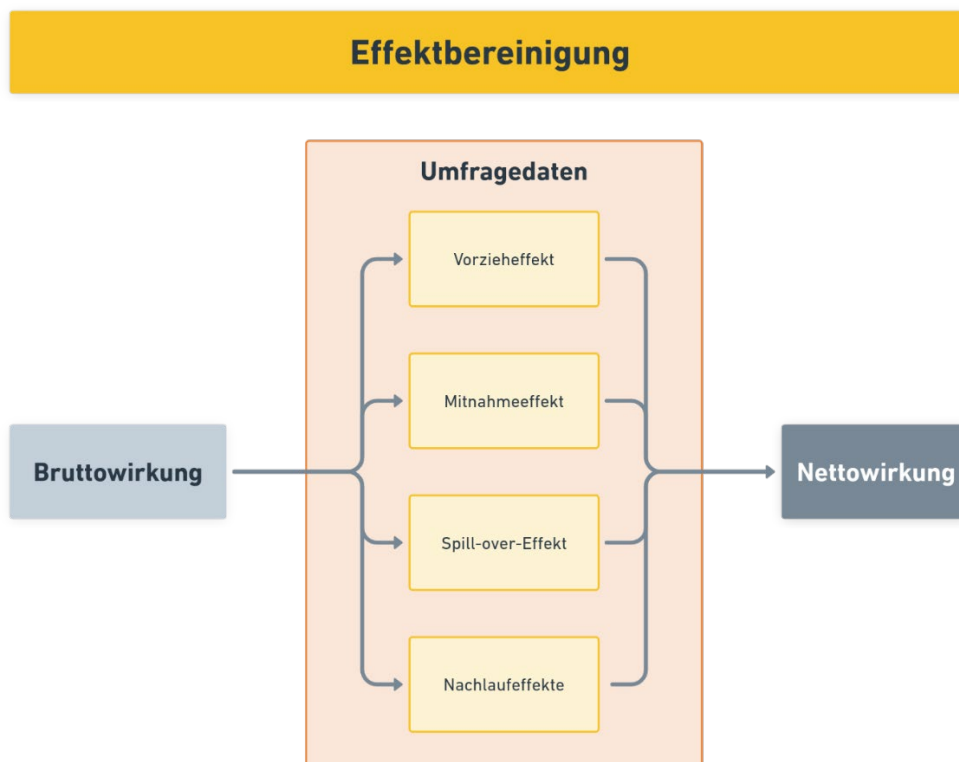


Abbildung 8: Effektbereinigung zur Berechnung der Nettoeffekte aus dem Brutto. Quelle: Eigene Darstellung

In der Praxis (bei veröffentlichten Bewertungen) wird der Bruttoeffekt verwendet, nicht die Nettoanpassung. Die Berechnungen, die mit der Verwendung von Brutto-/Nettoeffekten verbunden sind, sind in den Leitfäden so formuliert, dass beide Werte verwendet werden können (Abbildung 9). Je nach betrachteter Gruppe kann die Fehlermarge bis zu 20 % betragen, für die Gesamtgruppe bis zu 5 % [Interview 7]. Das bedeutet, dass bei der Ermittlung des Nettoeffekts mit erheblichen Unsicherheiten zu rechnen ist. In der untersuchten Stichprobe der veröffentlichten Evaluationen wurden in allen Fällen die Bruttoeffekte verwendet.

In einer Evaluierung der KfW 433:

Wie bereits dargelegt, können die Auswirkungen ermittelt werden. Aus Sicht des Evaluationsteams erscheint es jedoch nicht sinnvoll, die Wirkungen für die Ermittlung eines Nettoeffekts zu verwenden. Sie werden daher im Folgenden beschrieben und diskutiert, fließen aber nicht in die Wirkungsermittlung ein. (KfW-433, S. 84)

In der Ex-post-Evaluierung der EBS Arbeitsgruppe - Abschnitt 5.1.1 - Ansatz zur Anpassung der Auswirkungen:

Vor diesem Hintergrund stellt die Wirkungsermittlung mittels Befragungen eine Annäherung dar, deren Aussagekraft insbesondere im Hinblick auf die differenzierte Zuordnung zu einzelnen Programmen/Akteursgruppen nicht überinterpretiert werden sollte. Daher werden im Evaluationsbericht immer sowohl Brutto- als auch Nettowerte angegeben. (evaluation-kfw-foerderprogramme-ebs-wg-2018-2021, s. 61 - 62)

Fördereffizienz

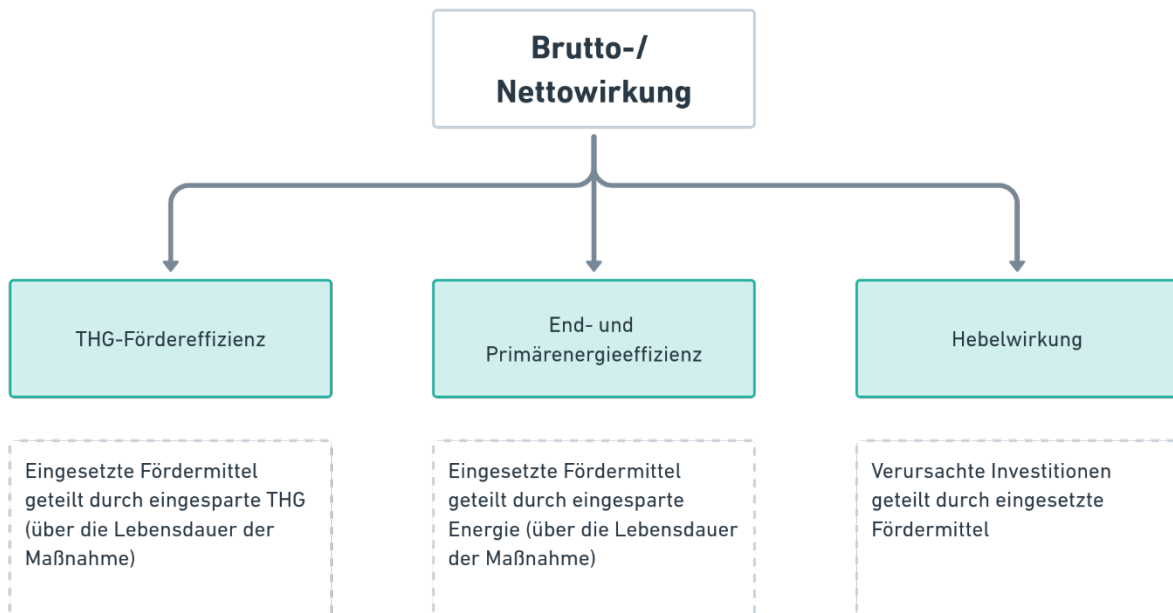


Abbildung 9: Verwendung von Brutto-/Nettoeffekten in Methodikleitfäden. Quelle: Eigene Darstellung

5.4.3. Interaktionseffekte

In den Methodikleitfäden werden Interaktionseffekte auf der Bündel- und nicht auf der Einzelmaßnahmenebene hervorgehoben. Sowohl in dem Ex-ante- als auch in dem Ex-post-Leitfaden wird die Bedeutung der Berücksichtigung von Interaktionseffekten auf Bündel- und nicht auf Einzelmaßnahmenebene anerkannt (Ex-post-Leitfaden, S. 99-100; Ex-ante-Leitfaden, S. 11-13).

Bei Ex-post-Evaluierungen besteht die Hauptmotivation für die Berücksichtigung von Wechselwirkungen darin, eine Doppelzählung von Energiespareffekten zu vermeiden. Im Zusammenhang mit Ex-post-Evaluierungen wird betont, dass Wechselwirkungen zwischen Energieeffizienzmaßnahmen, insbesondere innerhalb von Maßnahmenpaketen, untersucht werden sollten, um Doppelzählungen von Effekten zu vermeiden (Ex-post-Leitfaden, S. 99). Die Leitfäden schlagen verschiedene Ansätze vor, um diese Wechselwirkungen zu untersuchen, darunter die direkte deskriptive Bewertung, die analytische Bewertung oder die Erweiterung des Kreises der Unternehmen. Instrumentenfaktoren und

Interaktionsmatrizen können genutzt werden, um die Wechselwirkungen zwischen verschiedenen Instrumentenkombinationen transparent zu beschreiben.

Die Ex-ante-Leitfaden betonen auch die Notwendigkeit, Wechselwirkungen zwischen den Instrumenten zu berücksichtigen. In ähnlicher Weise wird in dem Leitfaden für Ex-ante-Evaluierungen die Bedeutung der Berücksichtigung von Interaktionseffekten innerhalb von Maßnahmenpaketen hervorgehoben (Ex-ante-Leitfaden, S. 11-13). Es werden verschiedene Methoden zur Bereinigung dieser Wechselwirkungen vorgeschlagen, mit oder ohne integrierte Modellierung. Diese Methoden können die Verwendung von Instrumentenfaktoren, Interaktionsmatrizen oder integrierten Modellen beinhalten. Integrierte Modellierungsrahmen erfassen implizit Wechselwirkungen zwischen einzelnen Maßnahmen und ermöglichen die Rückverteilung des Einspareffekts eines Maßnahmenpakets auf einzelne Maßnahmen durch Techniken wie Ranking, lineare Skalierung oder Skalierung mit Instrumentenfaktoren oder einer Instrumentenmatrix.

In der Praxis hielten sich die veröffentlichten Ex-post-Evaluierungen bei der Berücksichtigung von Interaktionseffekten nicht an die Leitfäden. Es ist jedoch anzumerken, dass die beiden Ex-post-Evaluierungen, die zu Förderprogrammen für energieeffiziente Gebäude in Deutschland durchgeführt wurden, insbesondere KfW 433 und KfW EBS-WG, Interaktionseffekte auf der Bündelungsebene berücksichtigt haben. Allerdings wurden bei keiner der beiden Evaluierungen die Empfehlung des Leitfadens zur Berücksichtigung dieser Wechselwirkungen explizit angewandt. Die KfW-433-Evaluierung untersuchte zwar die Inanspruchnahme anderer Förderprogramme neben der KfW-433, wandte aber nicht die empfohlenen Ansätze zur Berücksichtigung von Wechselwirkungen an (KfW-433, S. 108-109). Auch die KfW-EBS-WG-Evaluierung untersuchte die Bekanntheit und Nutzung anderer Förderprogramme, nutzte aber nicht die empfohlenen Ansätze zur Anpassung von Wechselwirkungen (KfW-EBS-WG-2018-2021, S. 76).

Aus unserer Stichprobe von Evaluierungen, die nach der Veröffentlichung der Leitfäden durchgeführt wurden, schließen wir, dass eine weitere Standardisierung für Interaktionseffekte erforderlich ist. Auch wenn unsere Stichprobe aufgrund begrenzter öffentlich verfügbarer Daten klein ist, wurden die Leitfäden in keiner dieser Evaluierungen explizit angewandt, um Interaktionseffekte zu berücksichtigen. Dies deutet darauf hin, dass eine weitere methodische Standardisierung hilfreich wäre.

5.4.4. Allgemeine Empfehlungen

Transparenz und Verlässlichkeit sind wichtige Aspekte bei den Evaluierungsmethoden.

Gegenwärtig konzentrieren sich die Evaluierungen in erster Linie auf statische Effekte und sollten auch die Auswirkungen der Maßnahmen über die gesamte Lebensdauer berücksichtigen, um eine umfassendere Bewertung zu ermöglichen. Die Replizierbarkeit ist von entscheidender Bedeutung, um die Glaubwürdigkeit und Robustheit der Evaluierungsergebnisse zu gewährleisten. Es sind weitere Anstrengungen erforderlich, um die Standardisierung der Evaluierungsverfahren zu verbessern und eine konsistente Replikation von Studien zu ermöglichen. Empfehlenswert sind Sensitivitätsanalysen, insbesondere dann, wenn es Unsicherheiten bei den Parametern gibt. Durch die Variation dieser Parameter hilft die Sensitivitätsanalyse, ihre potenziellen Auswirkungen auf die Evaluierungsergebnisse zu verstehen, und liefert verlässlichere Daten für Entscheidungsprozesse.

Die Einführung einer Dokumentation zur Offenlegung von Daten für beauftragte Evaluatoren würde die Transparenz und Reproduzierbarkeit deutlich verbessern.

Es ist derzeit nicht transparent, wie die Daten korrigiert werden, um fehlende Werte oder Daten, die subjektiv als „offensichtlich falsch“ angesehen werden, zu berücksichtigen. Um diese Probleme zu verbessern, sollten unterstützende Dokumente zur Offenlegung von Daten eingeführt werden, die vom Berater zusammen mit der politischen Bewertung beantwortet werden. Diese Daten sind oft als nicht-öffentliches Dokument verfügbar (müssen manchmal auch von der Revisionsabteilung aufgrund interner Unternehmensanforderungen zur Verfügung gehalten werden) [Interview 7]. Sie sollten jedoch öffentlich zugänglich gemacht werden. Ohne eine breite Zugänglichkeit einer solchen Datenerklärung/Datendokumentation ist eine Replikation derzeit nicht möglich.

Um Akzeptanz, Adoptionsraten und Zusätzlichkeit besser beurteilen zu können, werden alternative Methoden benötigt, um eine kontrafaktische Gruppe der breiteren Bevölkerung zu bilden. Während die Bundesbehörden, die die Subventionen verwalten, Daten von den Empfängern durch Selbstauskünfte erheben können, gibt es keine solchen Verpflichtungen für Gebäude, die keine Subventionen erhalten. Dies führt zu einer Verzerrung der Selbstauskunft, so dass die erhobene Stichprobe nicht repräsentativ für die Gesamtbevölkerung ist.

Eine Bewertung ist in der Regel voreingenommen. Wir wissen nur über diejenigen Bescheid, die Zuschüsse beantragen. Wir wissen nichts über diejenigen, die ähnliche Dinge ohne Zuschüsse oder ohne Finanzierung tun. [Interview 7]

Die Durchführung nationaler Umfragen kann helfen, eine repräsentativere Kontrollgruppe zu bilden. Prognos führt derzeit eine Umfrage zur Wirksamkeit des BEG⁵ durch. Diese sind noch von den Rücklaufquoten abhängig und können eine Tendenz zu Bevölkerungsgruppen aufweisen, die umweltbewusster sind oder eher Maßnahmen ergreifen, da sie die Unterstützungsmechanismen bereits kennen. Eine höhere Beteiligung könnte durch die Aufnahme von Fragen in umfassendere nationale Erhebungen oder durch Anreize zur Teilnahme durch Werbeangebote oder Rabatte in Zusammenarbeit mit Energiedienstleistern erreicht werden. Andere Optionen für die Gestaltung von Kontrollgruppen und kontrafaktische Analysen werden in den Methodikleitfäden ebenfalls als Möglichkeit anerkannt, aber aufgrund von Bedenken hinsichtlich der Kosten und des Datenschutzes häufig nicht angewendet. Unabhängige Forschungsarbeiten, wie das Wohnen-Wärme-Panell (Frondelet al. 2021), haben begonnen, diese Lücke zu schließen.

⁵ <https://www.prognos.com/de/FAQ-BEG>

6. Nutzung von Evaluierungen und Auswirkungen auf den politischen Prozess

In diesem Abschnitt werden die wichtigsten Verwendungszwecke von Evaluierungsergebnissen im politischen Prozess in Deutschland beschrieben. Die Abschnitte beziehen sich auf vier Hauptthemen: (i) analytische Beiträge zu Ex-ante-Prognosen; (ii) Motivation, Beauftragung von Evaluierungen; (iii) Zeitplanung von Evaluierungen und informelle Beratung; und (iv) Verbreitung von Evaluierungsergebnissen und Transparenz.

6.1. Analytische Beiträge zur Ex-ante-Prognose

Ex-post-Evaluierungen werden als Grundlage für Ex-ante-Evaluierungen benötigt. Ex-post-Evaluierungen sind ein entscheidender Aspekt der Governance, da sie den Nachweis erbringen, ob die politischen Maßnahmen wirksam sind, und die Grundlage für die Ex-ante-Evaluierungen bilden. Verlässliche und umfassende Ex-post-Evaluierungen sind zumindest für die wichtigsten Programme erforderlich, um gute Ex-ante-Evaluierungen vornehmen zu können [Interview 8].

Wenn Sie zum Beispiel ein Wohnbauförderungsprogramm haben, dann nehmen wir die Ex-post-Evaluierung und verwenden die wichtigsten relevanten Indikatoren für dieses Wirkungsmodell auch für die Ex-ante-Evaluierungen. [Interview 8]

Die Qualität und Zuverlässigkeit dieser Evaluierungen verbessert die Annahmen für Ex-ante-Wirkungsmodelle:

Annahmen werden durch empirisches Wissen gestützt, und in der Regel stützen wir uns bei Ex-ante-Evaluierungen auf die empirischen Zahlen aus den Ex-post-Evaluierungen. [Interview 8]

Darüber hinaus können Ungenauigkeiten bei der Ex-post-Evaluierung der Leistungsparameter von Instrumenten zu Ungenauigkeiten bei der Projektion künftiger Treibhausgas-minderungseffekte des geplanten Policy Mixes führen. Angesichts der zunehmenden Be-

deutung des Projektionsberichts als Auslöser für die im neuen Entwurf des Bundesklimagesetzes vorgesehene Reform des deutschen Klimamixes kann dies wichtige Auswirkungen auf die Steuerung der deutschen Klimapolitik haben.

Es gibt natürlich riesige Studien, auch sehr groß angelegte, zum Beispiel von Dena, die dann auch alle möglichen Szenarien für verschiedene Energieträger und natürlich auch sektorübergreifend beinhalten, das ist alles sehr hochkomplex. Aber die Daten zum Gebäudebestand sind regelmäßig ein Schwachpunkt. [Interview 4]

Oft mangelt es an Transparenz bei der Erstellung von Ex-ante-Evaluierungen. Die Annahmen bei der Modellierung sind weniger transparent oder nachvollziehbar als bei einigen der Ex-post-Evaluierungen. Dieser Mangel an Transparenz wirft die Frage auf, inwieweit wissenschaftliche Beiträge und Nachweise bei der Zielbildung und der Festlegung der Agenda berücksichtigt werden. Der begrenzte wissenschaftliche Beitrag in diesen Phasen kann die Wirksamkeit und Glaubwürdigkeit der politischen Entscheidungen beeinträchtigen.

Im Gegensatz zum Projektionsbericht, wo wir irgendwie als Vermittler beteiligt sind und den Prozess koordinieren, läuft dieser Quantifizierungsbericht immer ohne uns. Das heißt also, die Ergebnisse werden dann von denen verwendet, die auch die Ex-ante-Modellierung machen und wir kommen auch dazu, dass die Ergebnisse meist ein bisschen eingeschränkt sind. Wir sind aber nicht so direkt involviert [...] Also, wenn es um die Detailkurven geht, weiß ich nicht, wie detailliert das geplant ist. Da gibt es natürlich pro Modell welche, die das in unserem Auftrag machen [...] Experten, die das machen. [Interview 5]

In den Ex-ante-Leitfaden wird nicht zwischen verschiedenen Arten von Instrumenten unterschieden. Als ergänzendes Dokument erstrecken sich die Ex-ante-Leitfaden nicht auf die Unterscheidung zwischen verschiedenen Arten von Maßnahmen und beziehen sich nur auf die Ex-post-Evaluierungen.

Bei den Ex-ante-Evaluierungen werden z.B. Verwaltungskosten nicht berücksichtigt, und generell werden Governance-Kriterien außer Acht gelassen. Über die Verwaltungskosten hinaus werden die erforderlichen Kapazitäten, potenzielle Governance-Fehler (z.B. bei der Durchsetzung) und erforderliche Reformen (z.B. Konflikte mit bestehenden Gesetzen), die für eine effektive Umsetzung erforderlich sind, nicht berücksichtigt.

Dabei werden wichtige Evaluierungsaspekte, wie Engpässe bei der Verfügbarkeit von Handwerkern zur Durchführung von Renovierungsmaßnahmen und die Verfügbarkeit von Beratern, Inspektoren und Gutachtern übersehen. Ebenfalls bemerkenswert ist die fehlende Berücksichtigung eines robusten lokalen Inspektionssystems für Regulierungsmaßnahmen im Vollzug, die in diesem Sektor ein bekanntes Problem sind und die Wirksamkeit der Schätzungen für die Auswirkungen von Regulierungsmaßnahmen untergraben.

Die Lösung dieser Probleme würde einen Wechsel zu einem umfassenderen und inklusiveren Evaluierungsansatz erfordern, der eine breitere Palette von Kriterien als die fiskalischen Rechtfertigungen einbezieht. Dies würde eine größere Transparenz im Prozess der Zielbildung, einen verbesserten wissenschaftlichen Input und eine stärkere Einbeziehung der Interessengruppen erfordern, um eine effektivere und wirksamere Energie- und Klimapolitik zu fördern.

Das ist schwer zu beantworten. Wenn man sich die Landschaft der Instrumente und Programme im Gebäudesektor ansieht, so haben diese im letzten Jahr mit grundlegenden Veränderungen begonnen. Die meisten der neuen Instrumente und Programme sind also noch nicht richtig evaluiert. Die meisten Inputs für das Sofortprogramm wurden also aus einer Mischung von Ergebnissen älterer Evaluierungen und Teilen laufender Evaluierungen zusammengemischt und einfach modelliert oder mit Hilfe von Tools analysiert, um einen Überblick über mögliche Auswirkungen zu erhalten. Es ist also sehr reduziert. Ein Schlüsselmerkmal ist die Frage, wie man die so genannten weiteren Bemühungen erhält. Wenn Sie zum Beispiel einen Euro für Treibhausgasemissionen finanzieren, werden Treibhausgasemissionen oder Energie eingespart. Diese Informationen versuchen wir durch verschiedene Auswertungen zusammen zu stellen und auf den langfristigen Durchschnitt zu kommen. [Interview 7]

6.2. Motivation, Beauftragung und Nutzung von Evaluierungen zur Rekalibrierung der Maßnahmen

Die Motivation für die Beauftragung von Evaluierungen spielt eine entscheidende Rolle für deren Nutzung. In der Vergangenheit wurden Evaluierungen in erster Linie zur

Rechtfertigung von Steuerausgaben für Förderprogramme verwendet. Daher konzentrieren sich die Hauptkriterien, die bei diesen Bewertungen berücksichtigt werden, häufig auf die Bewertung von Energieeinsparungen und Kosteneffizienz. Die Einhaltung von Berichtspflichten durch die Ministerien scheint der Hauptnutzen von Evaluierungen zu sein. Die Erfüllung der Berichtspflichten ist entscheidend für die Zuweisung von Mitteln und die Gewährleistung der Rechenschaftspflicht bei der Umsetzung der Energie- und Klimapolitik.

Die Interviews deuten darauf hin, dass die derzeitige Governance von Evaluierungen die Beteiligung der Wissenschaft einschränkt und zu einer Voreingenommenheit führt.

Die Praxis, Evaluierungen in Auftrag zu geben, kann zu etablierten Beziehungen mit den wichtigsten am Evaluierungsprozess beteiligten Akteuren führen. Dies kann zu einem relativ geschlossenen Kreis von Analytinnen und begrenzten Möglichkeiten für externe Interessengruppen oder „Außenstehende“ führen, ihre Perspektiven oder ihr Fachwissen einzubringen [Interview 14]. Infolgedessen können Expertise und alternative Sichtweisen von externen Akteuren, z.B. aus der Wissenschaft, unberücksichtigt bleiben. Darüber hinaus werden die von den Bundesbehörden erhobenen Daten aufgrund des Datenschutzes nur den beauftragten Evaluatoren zur Verfügung gestellt. Sie sind nicht öffentlich verfügbar bzw. nicht für die Wissenschaft zugänglich, was die Möglichkeiten einer unabhängigen Analyse und Validierung einschränkt [Interview 7, 12, 14]. Diese fehlende gemeinsame Nutzung von Daten stellt ein Hindernis für die Transparenz dar und behindert das Potenzial für Forschung und die Validierung von Ergebnissen.

Mehrere Akteure innerhalb des Konsortiums der Hauptberater führen sowohl Ex-post-Evaluierungen als auch Ex-ante-Evaluierungen durch, was die Überprüfung der externen Validität potenziell einschränkt. Da die Ex-ante-Evaluierungen von denselben Akteuren durchgeführt werden, besteht die Gefahr eines ‚confirmation bias‘. Wären unabhängige Wissenschaftler und andere Akteure stärker beteiligt, könnten sie potentiell die Qualität, die Robustheit und den Umfang der Ex-post- und Ex-ante-Evaluierungen verbessern. Die in der KSG-Novelle vorgesehene stärkere Rolle des Projektionsberichts und die Rolle der ERK bei der Überprüfung des Berichts können die Transparenz und das Feedback verbessern. Wichtig ist, dass potenzielle Interessenkonflikte von Einzelpersonen ver-

mieden werden müssen. Dennoch sind Fortschritte bei den Evaluierungsverfahren zu verzeichnen, und die Standardisierung hat in letzter Zeit zugenommen. Im Jahr 2020 veröffentlichte das Konsortium der Hauptberater (Abschnitt 4.3) methodische Leitfäden, die ursprünglich für ein spezifisches Projekt gedacht waren [Interview 9]. Dieser Leitfaden wurde nun vom BMWK als Standardverfahren übernommen und muss bei allen Evaluierungen beachtet werden.

Der Datenzugang wird nur beauftragten Evaluatoren gewährt, was den Spielraum für unabhängige und akademische Forschung einschränkt. Der Zugang der Öffentlichkeit zu laufenden Arbeiten oder formativen Evaluierungen ist trotz der potenziellen Bedeutung dieser Berichte bemerkenswert eingeschränkt [Interview 6, 7]. Dieser Mangel an Transparenz behindert die Fähigkeit von Interessenvertretern, Forschenden und der Öffentlichkeit, den Fortschritt von Evaluierungen zu verfolgen und sich an informierten Diskussionen über die Ergebnisse zu beteiligen. Um dieses Problem anzugehen, ist es wichtig, der Transparenz im Evaluierungsprozess Priorität einzuräumen. Zur Förderung der Transparenz gehört es, der Wissenschaft Evaluierungsdaten zur Verfügung zu stellen, die Zusammenarbeit zu ermöglichen und eine externe Prüfung der Ergebnisse zu erleichtern. Indem externen Experten und Interessengruppen Zugang zum Evaluierungsprozess gewährt wird, können unterschiedliche Perspektiven einbezogen werden, was die Qualität und Glaubwürdigkeit der Bewertungen insgesamt erhöhen kann. Transparenz im Evaluierungsprozess ist für die Entwicklung evidenzbasierter Politiken von entscheidender Bedeutung. Sie ermöglicht eine fundierte Entscheidungsfindung und fördert das Vertrauen der Öffentlichkeit in die Ergebnisse der Evaluierung. Indem sichergestellt wird, dass Evaluierungsberichte, laufende Arbeiten und relevante Daten für ein breites Spektrum von Interessengruppen zugänglich sind, der Evaluierungsprozess transparenter und effektiver, was die Chance auf positive Veränderungen erhöht.

Das Fehlen von unabhängigen Analyse kann die Validität und den Umfang der Ex-post-Evaluierungen beeinträchtigen. Ohne externe Prüfung gibt es möglicherweise nur begrenzte Möglichkeiten für neue Ideen, die Verbesserung der Evaluierungspraxis oder die Kritik an bestehenden Ansätzen [Interview 12]. Das Fehlen eines externen Inputs und einer Validitätsprüfung kann die Möglichkeit einschränken, die Genauigkeit und Zuverlässigkeit der Evaluierungsergebnisse in Frage zu stellen, wodurch die Wirksamkeit und

Glaubwürdigkeit des Evaluierungsprozesses möglicherweise eingeschränkt wird [Interview 14].

Die von den Ministerien vorgegebene Spezifikation der Evaluierungsprozesse kann zu gewissen Verzerrungen führen, die sich auf die Ergebnisse der Evaluierungen auswirken. Die Hauptmotivation für die Durchführung der Evaluierung besteht darin, die Wirksamkeit von Programmen nachzuweisen und die Verwendung von Bundesmitteln zu rechtfertigen. Die in Auftrag gegebenen Evaluierungen können in ihrer Spezifikation einen ‚confirmation bias‘ verursachen, was bedeutet, dass die Ergebnisse wahrscheinlich keine radikalen Empfehlungen liefern oder systemische Fehler im Programm aufdecken [Interview 12]. Stattdessen konzentrieren sie sich eher auf kleine technische Vorschläge und schränken den Spielraum für die Analyse transformative Veränderungen ein. Solche Änderungen hängen mit dem begrenzten Umfang der bewerteten Kriterien (Abschnitt 6.2.) und den Programmtypen zusammen, die bewertet werden (Abschnitt 6.3.).

Die Rolle von Evaluierungen scheint während der vorherigen Regierungen einen relativ geringen Einfluss auf politische Entscheidungen gehabt zu haben. Aus unseren Interviews geht hervor, dass, wenn die Evaluierungen auf Probleme mit den Programmen hinwiesen (z.B. darauf, dass sich die Steuerausgaben nicht auf neue Gebäude oder auf sehr hohe Renovierungsstandards konzentrieren sollten (z.B. Standard KfW 45), die Empfehlungen nicht zu Reformen führten [Interview 6]. Die Häufung von politischen Rückmeldungen, dass das Programm nicht gut funktionierte, reichte nicht aus, um Reformen zu motivieren. Die relativ geringe Bedeutung und Nutzung früherer Evaluierungen könnte dazu beitragen, den relativ begrenzten Umfang zu erklären, den sie berücksichtigen. Eine verstärkte Rechenschaftspflicht durch die Umsetzung der KSG könnte dazu führen, dass Evaluierungen in Zukunft stärker genutzt werden.

Vieles von dem, was in der Wissenschaft und auch in den Instituten, die an den Themen arbeiten, eigentlich Konsens war, war in früheren Bundesregierungen in dieser Form politisch einfach nicht durchsetzbar. [Interview 6]

Die Auswirkungen von Ex-post-Evaluierungen auf die Entscheidungsfindung beschränken sich auf schrittweise Änderungen am laufenden Programm. Die in Auftrag gegebenen Evaluierungen dürften keine radikaleren Empfehlungen hervorbringen. Sie können Aufschluss darüber geben, ob die Maßnahme funktioniert oder nicht, und dienen dazu,

die Kosten des Programms zu rechtfertigen. Wie in Abschnitt 6.2 dargelegt, hängt die Zuverlässigkeit dieser Evaluierungen in hohem Maße von den vorgenommenen Anpassungen für die energetische Gebäudebilanzierung und den Annahmen für den ursprünglichen Energieverbrauch des Gebäudes vor den Maßnahmen ab. Derzeit ist nicht ausreichend transparent, wie diese Anpassungen vorgenommen werden. Externe Validierungstests deuten darauf hin, dass diese überschätzt werden, was sich sowohl auf die Wirksamkeit der aktuellen Maßnahmen als auch auf die Kosteneffizienz auswirkt. Mögliche Überschätzungen der Wirksamkeit der derzeitigen Programme schränken die Aussichten auf Evaluierungen ein, die zu radikaleren Reformen führen könnten.

Evaluierungen und Bewertungen spielen zwar eine Rolle, aber sie wirken sich in der Regel eher auf kleinere technische Änderungen als auf größere systemische Reformen aus. Im Bereich bedeutender politischer Veränderungen verlassen sich die politischen Akteure oft eher auf politisch relevante Themen oder Intuitionen als auf umfassende Evaluierungen. In manchen Fällen suchen sie selektiv nach Evaluierungen, die ihren vorgefassten Meinungen oder gewünschten Ergebnissen entsprechen.

Also all die Anpassungen, die jetzt am BEG vorgenommen worden sind, auch ziemlich ad hoc und kurzfristig, was da ausschlaggebend war, weiß ich nicht, aber letztlich in der wissenschaftlichen Debatte, die Hinweise kommen ja schon sehr, sehr lange, die Mittel auf die Baubereiche zu konzentrieren, wo sie wirklich gebraucht werden [...] ob das durch diese Evaluierungen und Studien oder durch einzelne Leute in den Ministerien getrieben worden ist, kann ich nicht beurteilen, aber letztlich sind einige Dinge umgesetzt worden, die in den Evaluierungen und Studien seit vielen, vielen Jahren immer wieder kritisiert worden sind.

[Interview 6]

6.3. Zeitplan für Evaluierungen und informelle Beratung

Die Zeitplanung ist ein entscheidender Faktor bei der Verbreitung von Evaluierungsergebnissen. Ein wichtiges Problem ist die zeitliche Verzögerung zwischen dem Forschungsabschluss, der Veröffentlichung der Ergebnisse und ihrer Integration in die politischen Diskussionen. Diese Verzögerung kann die rechtzeitige Nutzung von Evaluierungen für die Entscheidungsfindung und die Maßnahmenformulierung behindern.

Große Änderungen in den Programmen und die erforderliche Geschwindigkeit der Umstellung lassen keine Zeit für die Durchführung von Evaluierungen. Am problematischsten ist dies bei neuen Maßnahmen, für die es keine vorhandenen Erkenntnisse gibt, auf die man zurückgreifen könnte. Dies bedeutet, dass die Berater aufgefordert werden, in sehr kurzen Zeiträumen Ratschläge zu erteilen, was eine solide Evaluierungsarbeit nicht zulässt [Interview 9].

Im Moment ist das Tempo der Änderungen und Ergänzungen der Programme so hoch, dass wir manchmal, ich meine vor ein paar Wochen, an Sitzungen teilnehmen, bei denen wir, bevor wir auch nur eine kleine Bewertung vorgenommen haben, Empfehlungen abgeben sollen, wie die Programme geändert werden sollten. Ich meine, aber das ist wirklich die Situation im Moment, das Tempo und die Geschwindigkeit der erforderlichen Veränderungen [Interview 9].

Verzögerungen bei der Vorlage von Nachweisen führen dazu, dass die Bereitstellung von Informationen und die Konsultation in einer formalisierten oder kodifizierten Weise erfolgen. Der Rat von Experten, der auf früheren Erfahrungen beruht, kann bei der Gestaltung politischer Optionen offensichtlich eine wichtige Rolle spielen. Bestehende Beziehungen zu diesen Beratungsquellen haben oft Vorrang und schränken den Spielraum für die Berücksichtigung neuer Ideen und alternativer Optionen ein. Der Rückgriff auf informelle Kommunikation und der Einfluss von Beratern kann die Transparenz bei der Formulierung von Reformen reduzieren. Darüber hinaus erschwert der Mangel an öffentlich zugänglichen Daten die Möglichkeit einer unabhängigen Bewertung, die nicht direkt von den Ministerien in Auftrag gegeben wird. Diese Einschränkung führt zu einer Verzerrung bei der Auswahl der Akteure, die die Ministerien informieren können.

Meiner Meinung nach ist es besser, mit den Leuten zu sprechen, die diese Bewertungen vornehmen, denn sie können natürlich dieselbe Dynamik und den Markt und die Welt beobachten, und sie wissen auch, was vor sich geht, und sie haben eine wirklich gute Vorstellung. Und außerdem kennen sie die Daten, die natürlich nicht mehr ganz aktuell sind. Aber sie können diese beiden auch in einem aktuelleren Zeitrahmen zusammenbringen. Und das ist dann nichts, was in den Berichten steht, oder vielleicht in einigen dieser zusätzlichen Berichte, um die wir sie bitten, aber nicht in diesen Ex-post-Evaluierungen. [Interview 3]

Die Transparenz könnte durch partizipatorische Konsultationsprozesse erhöht werden, bei denen Erkenntnisse aus verschiedenen Wissensquellen eingeholt werden. Die Verbesserung der Transparenz durch die Veröffentlichung relevanter Daten würde unabhängige Bewertungen erleichtern und das Spektrum der bei der Politikformulierung berücksichtigten Informationen erweitern. Die Förderung der Vielfalt an Fachwissen und die Unterstützung eines integrativen Entscheidungsprozesses können ebenfalls dazu beitragen, unzulässige Einflussnahme zu vermeiden und einen robusteren und innovativeren Ansatz bei der Politikentwicklung zu fördern.

6.4. Verbreitung der Bewertungen und Transparenz

Eine explizitere und genauere Spezifikation der Anforderungen bei der Vergabe von Evaluationsaufträgen könnte die Wirksamkeit in der Beratung erhöhen. Der Mangel an Spezifität bei der Zusammensetzung von Evaluierungsteams behindert den Nutzen der Berichterstattung. Die interne Koordinierung zwischen den Mitgliedern der Beraterteams, die für die Durchführung von Evaluierungen verantwortlich sind, ist oft unzureichend, und es besteht häufig eine Trennung zum Auftraggeber [Interview 12]. Dieser Mangel an Koordination führt zu technischen Empfehlungen, die nur schwer in umsetzbare Strategien oder Maßnahmen umgewandelt werden können.

Um diese Probleme anzugehen, ist es wichtig, Vielfalt in die Evaluierungsteams einzubringen. Diese Vielfalt sollte nicht nur disziplinäres Fachwissen umfassen, sondern auch Personen mit unterschiedlichen Perspektiven und Hintergründen einbeziehen. Durch die Einbeziehung eines breiteren Spektrums von Experten in den Evaluierungsprozess kann die Kommunikation der Ergebnisse verbessert werden, wodurch sichergestellt wird, dass die technischen Empfehlungen für die politischen Entscheidungsträger leichter zugänglich und umsetzbar sind. Die Einführung von Diversitätsanforderungen in den Ausschreibungsprozess für Evaluierungsteams kann dazu beitragen, die Einschränkungen bei der Berichterstattung zu überwinden und die Qualität und den Nutzen von Evaluierungen zu verbessern.

Einige interviewte Stakeholder haben die Notwendigkeit einer stärkeren Einbeziehung und Konsultation in der Phase der Ziel- und Strategieformulierung geäußert. Einige Berater äußerten die Besorgnis, dass die Ziele ohne den Beitrag eines breiten Spektrums an Fachwissen festgelegt wurden [Interview 7, 11]. Ebenso hat die unzureichende Koordinierung zwischen den verschiedenen Verordnungen dazu geführt, dass für die einzelnen Programme unterschiedliche Ziele festgelegt wurden, was zu unterschiedlichen Anforderungen an die Berichterstattung geführt hat und die Transparenz und Verbreitung erschwert [Interview 7]. Diese Forderung nach einer stärkeren Einbindung und Mitwirkung einer Vielzahl von Interessengruppen deutet auf den Wunsch nach einem integrativeren und partizipativen Ansatz bei der Gestaltung von Energie- und Klimazielen hin. Dies könnte auch dazu beitragen, die Kohärenz in den verschiedenen Zielen und Strategien zu verbessern. Eine solche Einbeziehung kann dazu beitragen, dass die Ziele ausreichend ehrgeizig und gleichzeitig wissenschaftlich fundiert sind.

7. Die wichtigsten Reformoptionen

Die wichtigsten Reformoptionen werden nun für jeden Abschnitt skizziert und in entsprechenden Tabellen zusammengefasst.

7.1. Institutionelle Arrangements und Koordinierung

Eine Verbesserung der Koordinierungsmechanismen zwischen Bundesbehörden und Ministerien würde die Rechenschaftspflicht und die Zugänglichkeit der Daten verbessern. Es ist ratsam, strengere Maßnahmen zur Gewährleistung der Qualität der aufgezeichneten Daten einzuführen. Dies erfordert eine verstärkte Koordinierung und Zusammenarbeit zwischen Ministerien und Behörden. Die Einführung eines gemeinsamen Verhaltenskodexes für Behörden kann die Rechenschaftspflicht durch die Festlegung standardisierter Richtlinien fördern. Schlüsselaspekte, wie die Digitalisierung der erfassten Daten, sollten in diesem Rahmen vorrangig behandelt werden. Wenn die Daten öffentlich zugänglich gemacht und einer externen Überprüfung und Validierung unterzogen werden, entweder durch staatliche Stellen oder externe Kapazitäten, werden Transparenz und Rechenschaftspflicht weiter gestärkt. Durch diese Maßnahmen kann die Regierung eine Kultur der Rechenschaftspflicht fördern, die Zugänglichkeit der Daten verbessern und effizientere Entscheidungsprozesse ermöglichen.

Die Regierung kann eine größere Vielfalt in den Berater- und Evaluatorengruppen durch eine verstärkte Zusammenarbeit mit der breiteren wissenschaftlichen Gemeinschaft fördern. Um die Transparenz und die externe Kontrolle zu verbessern, könnte die Basis des Fachwissens der Berater und Evaluatoren verbreitert werden. Dies kann erreicht werden, indem die Transparenz von Daten und Annahmen erhöht wird und Daten öffentlich zugänglich gemacht werden, um externen Sachverstand und wissenschaftliche Forschung zu ermöglichen und so das politische Teilsystem zu öffnen. Durch den Zugang zu den Daten wird die Transparenz erhöht, was eine Rechenschaftspflicht ermöglicht und externen Expertinnen und Experten und der Zivilgesellschaft die Möglichkeit gibt, die Ergebnisse zu überprüfen. Die Einbeziehung von externem Fachwissen bringt unterschiedliche Perspektiven und unabhängige Beurteilungen mit sich und stärkt die Bewertungen. Dies könnte besonders wertvoll für Themenbereiche sein, die erst seit kurzem in die Evaluierungspraxis einbezogen werden, wie zum Beispiel Verteilungseffekte. Die Förderung

wissenschaftlicher Forschung fördert die evidenzbasierte Entscheidungsfindung und die Zusammenarbeit mit der Wissenschaft und erleichtert auch formale Evaluierungsberichte, die sich auf neu entwickelte Methoden stützen können. Dies könnte Teil der Ausschreibungsanforderungen sein, während die Forderung nach Transparenz der Daten und Annahmen die Zuverlässigkeit gewährleistet und eine externe Validierung ermöglicht.

Koordinierung zwischen Bund und Ländern. Die Verbesserung der Koordinierung zwischen Bund und Ländern beinhaltet mehrere Schlüsselreformen, die zu einer effektiveren Zusammenarbeit und Umsetzung der Maßnahmen beitragen können. Erstens wird empfohlen, Rechenschaftspflichten der Länder zu stärken. Dies kann durch die Einrichtung klarer Mechanismen für die Berichterstattung und Überwachung der Fortschritte, die Festlegung von Leistungszielen und die Gewährleistung von Transparenz in den Entscheidungsprozessen der Länder erreicht werden. Indem die Länder für ihre Maßnahmen und Ergebnisse zur Rechenschaft gezogen werden, kann die Koordinierung zwischen der Bundesregierung und den Ländern verstärkt werden. Zweitens ist die Förderung des Datenaustauschs zwischen Bund und Ländern eine wesentliche Voraussetzung für eine wirksame Koordinierung. Dazu gehört die Einrichtung von Mechanismen und Plattformen für den Austausch von Informationen, Daten und bewährten Verfahren. Die gemeinsame Nutzung relevanter Daten und Erkenntnisse kann eine evidenzbasierte Entscheidungsfindung erleichtern, die Kohärenz der Maßnahmen verbessern und eine bessere Koordinierung der Bemühungen von Bund und Ländern ermöglichen. Schließlich ist eine Aufstockung der Mittel für die Aufsichtsbehörden der Länder mit Unterstützung der Bundesregierung von entscheidender Bedeutung, um eine wirksame Aufsicht und Durchsetzung der Maßnahmen zu gewährleisten. Eine angemessene Finanzierung kann die Kapazität der Aufsichtsbehörden zur Durchführung von Überwachungs-, Einhaltung- und Durchsetzungsmaßnahmen stärken. Dies wiederum trägt dazu bei, die Integrität von Maßnahmen und Vorschriften zu wahren, die Rechenschaftspflicht zu verbessern und die erfolgreiche Umsetzung gemeinsamer Initiativen von Bund und Ländern zu unterstützen.

Aktuelle Probleme mit der institutionellen Konfiguration und mögliche Reformationen		
Problem	Auswirkungen	Mögliche Lösungen
Begrenzte Koordinierung und Rechenschaftspflicht bei der Aktenführung in den Bundesministerien	Die aufgezeichnete variable Datenqualität erschwert die Bewertung und erfordert von den Bewertern, Fehler zu korrigieren und Anpassungen vorzunehmen.	<ul style="list-style-type: none"> Formelle Prüfung der föderalen Agenturen in Bezug auf die operationellen Verfahren durch die für die Leitung zuständigen Ministerien. Interne Berichtsanforderungen.
Mangelnde vertikale Koordinierung und mangelnder Datenaustausch zwischen Bund und Kommunen führen zu einer Rechenschaftslücke der Länder bei der Durchsetzung des GEG.	Die mangelnde Glaubwürdigkeit der Durchsetzung der GEG mindert die Wirksamkeit.	<ul style="list-style-type: none"> Fragen des Datenschutzes und der gemeinsamen Nutzung von Daten durch Bund und Länder sollten geklärt werden.
Die Durchsetzung bzw. Einhaltung der Vorschriften wird an die Landkreise delegiert, denen es in der Regel an Mitteln und Fachwissen fehlt, um diese Aufgabe zu erfüllen.	Die Landkreise sind nicht in der Lage, ihre Rolle als Aufsichts- und Vollzugsorgan wirksam zu erfüllen.	<ul style="list-style-type: none"> Eine neue Bundesbehörde oder eine Erweiterung des Kompetenzbereichs einer bestehenden Behörde, die sich mit der Durchsetzung von Rechtsvorschriften befasst. Diese Behörde sollte die Durchsetzungs- und Inspektionsverfahren standardisieren. Es sollten Mittel für die Entwicklung von Bewertungsmethoden bereitgestellt werden. Zur Erfüllung dieser Aufgabe sollten regionale Unterabteilungen eingerichtet werden, die direkt von der zentral koordinierenden Aufsichtsbehörde geleitet werden. Die Finanzmittel und die Ausbildung für die lokalen Inspektoren sollten aufgestockt werden.
Relativ geschlossene Gruppe und stabiles Konsortium von Beratern.	Begrenzt die Innovation von Evaluierungsmethoden und -praktiken.	<ul style="list-style-type: none"> Erweiterung des Beraterkreises. Förderung der Zusammenarbeit mit anderen Fachleuten.

Tabelle 9: Wichtigste Reformationen zur Verbesserung des institutionellen Aufbaus.

7.2. Umfang der bewerteten Programme und Evaluationskriterien

Das GEG wird derzeit nicht evaluiert, was eine erhebliche Lücke in der Steuerung des Wohnungsbausektors darstellt. Das Fehlen von Regulierungsinstrumenten in Bewertungs- und Evaluierungsprozessen kann darauf zurückgeführt werden, dass Regulierungsstandards auf regionaler (Länder-)Ebene verwaltet werden, zusammen mit Datenschutzgesetzen, die den Informationsaustausch zwischen lokalen Behörden und Bundesministerien einschränken, wie bereits erwähnt. Infolgedessen mangelt es an der Aggregation und Einbeziehung regionaler Daten in Bewertungen und Beurteilungen. Die Lösung dieser Probleme sollte eine Priorität sein, auch wenn die Koordinierung zwischen

lokalen und föderalen Regierungen weiterhin schwierig ist. Eine mögliche Lösung könnte darin bestehen, Beratungsunternehmen Zugang zu lokalen Aufzeichnungen zu gewähren und sie in die Lage zu versetzen, die Daten für eine umfassende Evaluierung zu aggregieren. Die jüngste Einführung von Effizienzstandards kann dazu beitragen, die Berichterstattung über die Einhaltung von Vorschriften zu verbessern. Es ist erwähnenswert, dass die Evaluierungs- und Beurteilungsmethoden nach der Einführung von Effizienzstandards in Deutschland im Jahr 2021 derzeit möglicherweise aktualisiert werden, obwohl aktualisierte methodische Leitfäden noch nicht öffentlich verfügbar sind.

Die Ausweitung des Umfangs aktueller Evaluierungen von Energieeffizienzmaßnahmen beinhaltet die Berücksichtigung von Verteilungseffekten und Governance-Herausforderungen. Es ist von entscheidender Bedeutung, die sozioökonomischen Auswirkungen dieser Maßnahmen zu berücksichtigen, um eine gerechte Verteilung der Vorteile zu gewährleisten und eine Verschärfung von Ungleichheiten zu vermeiden. Die jüngsten Evaluierungen sind ein Schritt in die richtige Richtung, da sie auch die Ausrichtung der BEG-Subventionen berücksichtigen. Derzeit zahlen sozialpolitische Programme für benachteiligte Bürgerinnen und Bürger, die staatliche Unterstützung erhalten und nicht arbeiten können, die Nebenkosten von staatlich geförderten Wohnungen. Die Kosten für diese Sozialprogramme werden nicht mit den Steuergeldern für die Förderprogramme zur Steigerung der Energieeffizienz verrechnet. Entscheidend ist, dass die Bewertung der Auswirkungen des GEG auf die Einkommensgruppen ebenfalls Priorität hat und, wenn möglich, als Ex-ante-Evaluierung durchgeführt wird, um so potenzielle negative Folgen abzumildern. Eine bessere Integration von Sozial- und Klimapolitik sollte auch zu einer besseren Berücksichtigung der Kostenwirksamkeit führen.

Governance-Herausforderungen spielen eine wichtige Rolle bei der Operationalisierung von Maßnahmen, werden aber derzeit bei Evaluierungen nicht berücksichtigt. Dies erfordert spezifische Verwaltungskapazitäten, einschließlich einer besseren Koordination zwischen den Finanzierungsstellen. Eine bessere Koordination kann die Effizienz und Effektivität des Vollzugs durch die Straffung von Prozessen und die Vermeidung von Doppelarbeit erhöhen. Die Bewältigung von Governance-Herausforderungen erfordert einen vielschichtigen Ansatz, der die Stärkung der Verwaltungskoordination, die Ver-

besserung der Durchsetzungskapazitäten und die Einführung robuster Überwachungsmechanismen umfasst. Auf diese Weise kann die Energieeffizienzpolitik wirksam umgesetzt werden, was eine gerechte Verteilung des Nutzens gewährleistet und sinnvolle Fortschritte auf dem Weg zu nachhaltigen und effizienten Energiesystemen ermöglicht.

<u>Aktuelle Probleme mit dem Anwendungsbereich und mögliche Reformoptionen</u>		
Problem	Auswirkungen	Mögliche Lösungen
GEG wird derzeit nicht bewertet.	<ul style="list-style-type: none"> Die NAPE-Berichterstattung und alle Ex-ante-Evaluierungen der Auswirkungen von Rechtsvorschriften auf den Bausektor beruhen nicht auf Ex-post-Nachweisen. Die Überarbeitung des GEG stützt sich nicht auf zuverlässige Ex-post-Evidenz. 	<ul style="list-style-type: none"> Einführung von Berichtspflichten. Standardisierte Bewertungsanforderungen und -verfahren für alle lokalen Behörden. Zugänglichmachung der in den Ländern erfassten Daten für eine koordinierende Evaluierungsberatung.
Steuerliche Anreize und BEHG nicht in hoher Auflösung bewertet.	<ul style="list-style-type: none"> Es ist nicht möglich, Änderungen beim Energieverbrauch und der Treibhausgasreduzierung genau auf Änderungen der Energiepreise zurückzuführen. 	<ul style="list-style-type: none"> Erhebung von Daten über den Energieverbrauch in höherer Auflösung. Verbesserte Gebäudedatenbank.
Umfang der Indikatoren ist begrenzt	<ul style="list-style-type: none"> Wichtige Informationen werden von den aktuellen Evaluierungen ausgeschlossen, was die Entscheidungsgrundlage einschränkt. 	<ul style="list-style-type: none"> Die Evaluierungen sollten mehr Indikatoren umfassen, insbesondere: Verteilungseffekte, Governance, dynamische Kostenwirksamkeit.
Die Evaluierungen beziehen sich nicht ausdrücklich auf Lücken/Herausforderungen bei der Governance oder auf spezifische Herausforderungen beim Datenzugang und der Datenqualität.	<ul style="list-style-type: none"> Die Regierung wird nicht ausdrücklich auf die notwendigen Reformen zur Verbesserung der Regierungsführung hingewiesen. 	<ul style="list-style-type: none"> Die Evaluierung sollte auf die Ermittlung von Governance-Herausforderungen ausgedehnt werden, einschließlich Kapazitäts-/Koordinations-/Durchführungslücken und spezifischer Probleme mit dem Zugang und der Qualität von Daten.

Tabelle 10: Wichtigste Reformoptionen zur Verbesserung des Umfangs der bewerteten Programme und der Bewertungsmaßstäbe.

7.3. Datenzugang und -qualität

Die Verbesserung des Zugangs zu und der Qualität von Daten stellt eine entscheidende Herausforderung dar, um die Zuverlässigkeit von Evaluierungen zu erhöhen. Die Bewältigung dieser Herausforderung erfordert Reformen der Datenschutzgesetze, um eine größere Verfügbarkeit und gemeinsame Nutzung von Daten zu ermöglichen. Während sensible Daten weiterhin geschützt werden müssen, sollten grundlegende Informationen, wie die Anzahl der Häuser, leichter zugänglich gemacht werden. Darüber hinaus ist die Digitalisierung von Daten von entscheidender Bedeutung, um mehr Transparenz und Rechenschaftspflicht zu erreichen, insbesondere für staatliche Fördereinrichtungen. Durch die Digitalisierung von Daten wird es einfacher, Informationen zu überprüfen und zu verifizieren, was zu einer höheren Qualität der Daten und solideren Evaluierungen führt.

Um den Zugang zu den relevanten Daten zu erleichtern, ist die Erstellung einer öffentlich zugänglichen Datenbank über den Gebäudebestand erforderlich. Diese Datenbank sollte Informationen wie die Art der Wohnungen, die vorhandenen Heizsysteme und Wärmeverluste der Bauteile (U-Werte) oder Effizienzklassen enthalten. Indem diese Daten Forschenden, politischen Entscheidungstragenden und der Öffentlichkeit zugänglich gemacht werden, fördern sie die Transparenz, ermöglichen eine faktengestützte Entscheidungsfindung und erleichtern die Überwachung und Bewertung von Energieeffizienzmaßnahmen. Durch die Reform der Datenschutzgesetze, die Digitalisierung von Daten und die Einrichtung einer öffentlich zugänglichen Datenbank mit Informationen zum Gebäudebestand, kann der Evaluierungsprozess von einem verbesserten Zugang zu zuverlässigen und umfassenden Daten profitieren. Dies wiederum stärkt die Zuverlässigkeit von Evaluierungen und erhöht die Wirksamkeit von Energieeffizienzmaßnahmen und -programmen.

Aktuelle Probleme mit dem Datenzugang/der Datenqualität und mögliche Reformoptionen		
Problem	Auswirkungen	Mögliche Lösungen
Die von den Bundesministerien erfassten Daten sind nur begrenzt aussagekräftig, da wichtige Faktoren wie die Zielorientierung (d. h. Verteilungseffekte) nicht berücksichtigt werden.	Begrenzt das, was solide bewertet werden kann, ohne dass es auf Erhebungen oder Näherungen beruht.	<ul style="list-style-type: none"> • Wichtige Informationen wie z. B. die Einkommensgruppe sollten angegeben werden. • Diese Informationen sollten in den BEG-Antrag aufgenommen werden, was auch die Ausrichtung verbessern würde, um die Zusätzlichkeit der Subventionen zu verringern.
Die Registerführung in den Bundesministerien speichert Daten mit den Namen der Antragsteller.	Die Daten fallen unter die Datenschutzgesetze für personenbezogene Informationen, die den Zugang und die gemeinsame Nutzung von Daten einschränken.	<ul style="list-style-type: none"> • Die Daten sollten anonymisiert erfasst werden. Einige identifizierende Merkmale, wie z. B. demografische Informationen, sollten aufgezeichnet werden, um die Extrapolation zu erleichtern. • Alternativ sollten diese Daten in der Nachbearbeitung anonymisiert oder pseudonymisiert werden.
Die Aktenführung in den föderalen Agenturen ist nicht digitalisiert.	Schränkt den Zugang zu den Daten ein und begrenzt die Möglichkeit von Qualitätskontrollen der Daten.	<ul style="list-style-type: none"> • Die Digitalisierung in allen Bundesbehörden sollte Priorität haben, um den Zugang zu verbessern, die Qualitätskontrolle der aufgezeichneten Daten und die Rechnungsprüfung zu ermöglichen.
Fehlen einer Gebäudedatenbank	Schränkt die Möglichkeit ein, genauere Bewertungen durchzuführen, indem zuverlässige Daten über den Energieverbrauch vor den Maßnahmen bereitgestellt werden.	<ul style="list-style-type: none"> • Änderung der Datenschutzgesetze, um die Verwendung von Zertifikaten zu ermöglichen. • Erhöhte Auflösung der Analyse von mehrstufigen Stichprobendaten.

Tabelle 11: Wichtige Reformoptionen zur Verbesserung des Datenzugangs und der Datenqualität.

7.4. Methoden

Monitoring der tatsächlichen Verbrauchswerte nach der Sanierung und Nutzung dieser Daten bei der Evaluierung. Monitoring ist für die Bewertung der Wirksamkeit von Energieeffizienzmaßnahmen unerlässlich. Sie sollte ein Monitoring in Echtzeit beinhalten, um ein sofortiges Feedback zu geben und rechtzeitige Anpassungen zu ermöglichen. Darüber hinaus bietet Monitoring über einen längeren Zeitraum, in der Regel 3-5 Jahre, Einblicke in die langfristigen Auswirkungen der Maßnahmen und hilft, Trends und verbesserungswürdige Bereiche zu erkennen. Die Verwendung von Daten über den Verbrauch vor der Umrüstung nur dann, wenn es sich um den tatsächlich gemessenen Verbrauch über einen Zeitraum von drei Jahren handelt, erhöht die Genauigkeit, stellt jedoch eine Herausforderung für die derzeitige Umsetzung der Maßnahmen dar, da es keine bestehende Datenbank gibt, auf die zurückgegriffen werden kann.

Die Standardisierung und Transparenz der Anpassungen und Annahmen im Evaluierungsprozess müssen weiter verbessert werden, um die Genauigkeit und Vergleichbarkeit zu erhöhen. Bei der Schätzung der Brutto- und Nettoeffekte besteht derzeit die Gefahr der Überschätzung und die verschiedenen Beratungsunternehmen wenden unterschiedliche Verfahren an. Eine Angleichung und Verfeinerung der Schätzmethoden kann zu präziseren Bewertungen führen. Um eine Standardisierung zu erreichen, ist eine bessere Koordinierung zwischen den Beratungsunternehmen erforderlich, die über den begrenzten Anwendungsbereich der bestehenden Leitfäden hinausgeht. Die Ministerien können eine zentrale koordinierende Rolle spielen, indem sie für einheitliche Evaluierungsverfahren und -methoden sorgen und die Auftragsfristen verlängern, um die Transparenz zu erhöhen. Die Ministerien müssen jedoch möglicherweise ihre Kapazitäten ausbauen, um ihr technisches Fachwissen über Evaluierungsmethoden zu erweitern. Schulungen, Ressourcen und Beratung können die Ministerien in die Lage versetzen, effektiv zur Standardisierung beizutragen. Die Förderung einer weiteren Standardisierung, die Erleichterung der Beratungskoordination und die Stärkung der technischen Kapazitäten der Ministerien werden zu soliden Evaluierungen, genauen Schätzungen der Auswirkungen und fundierten Entscheidungen für eine wirksame Maßnahmenumsetzung führen.

Aktuelle Probleme mit Methoden und mögliche Reformationen		
Problem	Auswirkungen	Mögliche Lösungen
Interaktionseffekte werden nur teilweise berücksichtigt und es fehlt an Transparenz und Standardisierung.	Wichtige Synergien oder notwendige Wechselwirkungen für die Wirksamkeit von Maßnahmen werden nicht ausreichend gemessen. Dies kann zu einer ineffizienten Planung und Verteilung von Ressourcen führen.	<ul style="list-style-type: none"> Die Methodik für das Zusammenwirken der politischen Instrumente sollte als Teil der Bottom-up-Evaluierungsmethodik weiterentwickelt werden. Weitere Standardisierung der Praktiken bei angewandten Bewertungen.
Begrenzter Einsatz von Kontrollgruppen-Design und kontrafaktischer Analyse: Dieser Ansatz ist als Option anerkannt, wird aber aufgrund von Bedenken hinsichtlich der Kosten und des Datenschutzes häufig nicht angewendet.	Schwierigkeiten bei der Bestimmung der Auswirkungen von Maßnahmen, der Zusätzlichkeit und der Auswirkungen von Marktmechanismen wie Steuern.	<ul style="list-style-type: none"> Es sollten Mittel für die Durchführung von Studien mit randomisierten Kontrollgruppen bereitgestellt werden. Dies könnte mit führenden wissenschaftlichen Instituten durchgeführt werden. Fragen des Datenschutzes müssten entweder im Rahmen allgemeinerer Reformen oder forschungsspezifischer Ausnahmen bzw. Befreiungen gelöst werden.

<p>In den Evaluierungen werden die Annahmen nicht klar dargelegt.</p>	<p>Geringere Transparenz, Reproduzierbarkeit und Zuverlässigkeit</p>	<ul style="list-style-type: none"> • Die Gewährleistung von Transparenz und Zuverlässigkeit der Bewertungsmethoden ist von entscheidender Bedeutung. Gegenwärtig konzentrieren sich die Evaluierungsverfahren in erster Linie auf statische Effekte, und es besteht die Notwendigkeit, die Auswirkungen von Maßnahmen über die gesamte Lebensdauer hinweg zu berücksichtigen. • Die Verbesserung der Replizierbarkeit von Bewertungen ist wichtig, um die Glaubwürdigkeit und Robustheit der Ergebnisse zu gewährleisten. Weitere Standardisierungsbemühungen sind notwendig.
---	--	---

Tabelle 12: Die wichtigsten Reformationen zur Verbesserung der Evaluationsmethoden.

7.5. Nutzung der Evaluierungen und Verbreitung

Konsistenz zwischen den verschiedenen Berichtsansforderungen. Die Sicherstellung der Konsistenz zwischen den verschiedenen Berichtsansforderungen ist für die Verbesserung der externen Transparenz und die Erleichterung der Übersetzung von Ergebnissen zwischen verschiedenen Programmen von wesentlicher Bedeutung. Unstimmigkeiten entstehen jedoch häufig aufgrund unterschiedlicher Ergebnisse und Parameter, die in verschiedenen Verordnungen festgelegt sind. Um dieses Problem zu lösen, ist eine stärkere Koordinierung zwischen den für diese Programme zuständigen Ministerien sowie die Festlegung gemeinsamer Ziele erforderlich. Das Vorhandensein unterschiedlicher Outputs und Parameter für die Berichterstattungsansforderungen kann zu Problemen beim Verständnis und Vergleich der Ergebnisse verschiedener Programme führen. Dieser Mangel an Konsistenz verringert die externe Transparenz, da es für die Beteiligten schwierig ist, die gemeldeten Ergebnisse effektiv zu interpretieren und zu vergleichen. Es ist wichtig, klare und standardisierte Berichtsansforderungen festzulegen, die Schlüsselindikatoren und Messgrößen umfassen, die für die Ziele der Programme relevant sind.

Um Konsistenz zu erreichen, sind Koordinierungsbemühungen zwischen den für diese Programme zuständigen Ministerien von entscheidender Bedeutung. Indem sie ihre Ziele aufeinander abstimmen und zusammenarbeiten, können die Ministerien die Berichtsansforderungen straffen und sicherstellen, dass dieselben Ergebnisse und Parameter in den verschiedenen Programmen einheitlich verwendet werden. Diese Koordinierung sollte

eine regelmäßige Kommunikation und Zusammenarbeit beinhalten, um die Berichtsrahmen zu harmonisieren und gemeinsame Richtlinien für die Datenerfassung, -analyse und -berichterstattung festzulegen. Darüber hinaus ist es wichtig, gemeinsame Ziele für alle Programme festzulegen. Wenn die Programme unterschiedliche Ziele verfolgen, wird es schwieriger, eine einheitliche Berichterstattung zu erreichen. Durch die Festlegung gemeinsamer Ziele und die Ausrichtung der Programme auf diese gemeinsamen Ziele wird es einfacher, einheitliche Berichtsrahmen zu erstellen. Dadurch können die Beteiligten die Ergebnisse der verschiedenen Programme besser vergleichen und bewerten.

Die Regierung kann – entsprechenden politischen Willen vorausgesetzt – die Reform der Evaluierungsverfahren sowie einen Ausbau der internen Kapazitäten zur Festlegung des Umfangs und der Methoden aktiv vorantreiben. Eine bessere Spezifikation der Bewertungsparameter würde den Wert der Ergebnisse erhöhen und bessere Prognosen und Ex-ante-Evaluierungen ermöglichen. Um die EU-Anforderungen zu erfüllen und eine bessere Bewertung und Vorhersage zu ermöglichen, sollte den Verteilungseffekten besondere Aufmerksamkeit geschenkt werden, um regressive Ergebnisse und politische Rückwirkungen zu vermeiden. Darüber hinaus ist es für eine umfassende Bewertung unerlässlich, den Umfang der Evaluierungen auf Fragen der Wirksamkeit der Governance, der Verwaltung und der Informationsanforderungen auszuweiten. Durch diese Schritte kann die Regierung sinnvolle Reformen vorantreiben und sicherstellen, dass die Evaluierungen wertvolle Erkenntnisse für eine fundierte Entscheidungsfindung und politische Verbesserungen liefern.

<u>Aktuelle Probleme bei der Nutzung/Verbreitung und mögliche Reformoptionen</u>		
Problem	Auswirkungen	Mögliche Lösungen
Ex-ante-Evaluierungen sind in Bezug auf Datenquellen und Nachweise nicht so transparent.	Für Ziele und Programme fehlt möglicherweise eine verlässliche Evidenzbasis.	<ul style="list-style-type: none"> Formalisierung der Offenlegung, Transparenz und Verwendung von Fakten bei der Entscheidungsfindung.
Inkohärente Ziele in den verschiedenen Strategien und Zielsetzungen führen zu unterschiedlichen Anforderungen an die Berichterstattung.	Die Ergebnisse der verschiedenen Programme werden auf der Grundlage unterschiedlicher Annahmen berechnet und unterschiedlich dargestellt, was ihre Verbreitung erschwert.	<ul style="list-style-type: none"> Einbeziehung einer größeren Vielfalt von Sektorspezialisten und Beratern in den Prozess der Ziel- und Strategieformulierung. Gewährleistung der Kohärenz zwischen den verschiedenen Berichtsanforderungen

Tabelle 13: Wichtigste Reformoptionen zur Verbesserung der Nutzung von Evaluierungen und der Verbreitung von Informationen.

7.6. Vereinbarkeit der wichtigsten Reformen mit der derzeitigen institutionellen Struktur

Die wichtigsten Reformoptionen werden im Hinblick auf ihre Vereinbarkeit mit der derzeitigen institutionellen Konfiguration dargestellt. Die Reformen werden im Hinblick auf ihre erwarteten Auswirkungen auf die Ermöglichung qualitativ besserer Evaluierungsprozesse dargestellt (Abbildung 10). Hochgradig kompatible Reformoptionen sind solche, die ohne wesentliche Änderung der eingebetteten formalen institutionellen Regeln (wie z.B. Datenschutzgesetze) umgesetzt werden könnten. Optionen mit derzeit geringer Kompatibilität mit dem institutionellen Gefüge, die jedoch ein hohes Wirkungspotenzial bieten, sollten dennoch als notwendig priorisiert werden, wobei zu berücksichtigen ist, dass Verfahrensregeln und der erforderliche politische Konsens die Geschwindigkeit, mit der diese eingebetteten Regeln reformiert werden könnten, beeinträchtigen können.



Abbildung 10: Reformoptionen, geordnet nach Auswirkungen und Vereinbarkeit mit der derzeitigen institutionellen Struktur. Quelle: Eigene Darstellung

8. Schlussfolgerungen

In diesem Bericht wird ein analytischer Ansatz entwickelt und verwendet, der integriert das institutionelle Setup und die Qualität und Nutzung von Evaluierungen, Monitoring und Berichterstattung analysiert. Durch diesen Ansatz wird ein umfassendes Verständnis der Gestaltung, Beauftragung und Durchführung von Evaluierungsprozessen erreicht. Dies geht über die bisherige Forschung hinaus, die sich in erster Linie auf die Qualität von Evaluierungsergebnissen oder auf staatliche Kapazitäten konzentriert, ohne diese beiden Aspekte miteinander zu verbinden.

Ungenauigkeiten bei der Evaluierung der Effekte von Instrumenten können erhebliche Auswirkungen auf die Projektion der künftigen Treibhausgasmindeffekte von Policy-Mixes haben. Es besteht die Gefahr, dass die derzeitigen Evaluierungen die THG-Minderungseffekte der Programme überschätzen. Dies ist auf zwei Hauptfaktoren zurückzuführen: Erstens werden die Einsparungen durch installierte BEG-Maßnahmen aufgrund von Prebound-Effekten möglicherweise überschätzt, d.h. der Energieverbrauch vor der Installation von Maßnahmen wird zu hoch angesetzt, was zu größeren erfassten Energieeinsparungen als in Wirklichkeit führt. Zweitens: Wenn die Vorschriften nicht wirksam kontrolliert und durchgesetzt werden, kann es sein, dass ein erheblicher Teil der erwarteten THG-Reduzierungen durch das BEG-Programm aufgrund von Umgehung und Nichteinhaltung nicht eintritt. Auch wenn es den Rahmen dieses Berichts sprengen würde, die kombinierten Auswirkungen dieser Faktoren auf die Fortschritte bei der Verringerung der Treibhausgasemissionen abzuschätzen, ist es plausibel, dass weitere Emissionsreduktionen in diesem Sektor durch unrealistische Erwartungen hinsichtlich der Wirksamkeit der derzeit eingesetzten Instrumente in den Ex-ante-Evaluierungsmodellen behindert werden könnten. Im schlimmsten Fall könnte das in den derzeitigen Minderungsstrategien verankerte Ambitionsniveau nicht ausreichen, um die Minderungsziele zu verwirklichen, da eine Leistungslücke zwischen der erwarteten Wirksamkeit und der Realität besteht.

Ein wichtiger Bereich, der Anlass zur Sorge gibt, ist das Fehlen einer soliden Evaluierung des Gebäudeenergiegesetzes (GEG). Die derzeitigen Evaluierungsprozesse für das GEG beschränken sich auf einige Regulierungsmaßnahmen, die in der NAPE-Berichter-

stattung enthalten sind. Es fehlt jedoch an verlässlichen Daten für viele notwendige Kriterien, um deren Auswirkungen gründlich zu bewerten. Darüber hinaus untergraben die mangelnde Durchsetzung und die weit verbreitete Nichteinhaltung der Vorschriften im Bausektor die Wirksamkeit der bestehenden Regelungen. Die Evaluierung und der Ausbau der Verwaltungskapazitäten zur wirksamen Verwaltung und Durchsetzung von Regulierungsmaßnahmen ist eine wichtige Priorität. Datenbeschränkungen und Schwierigkeiten beim Zugang zu regionalen Daten stellen ein großes Problem für die Rechenschaftspflicht dar und untergraben die Glaubwürdigkeit der für die Durchsetzung der Vorschriften zuständigen Aufsichtsbehörden. Dieser Mangel an Transparenz und Rechenschaftspflicht erschwert zudem die Bewertung der in den Vorschriften festgelegten Energieeffizienzziele. Die Beseitigung dieser Fälle, die mit erheblichen Unsicherheiten behaftet sind, ist eine der wichtigsten Prioritäten.

Der derzeitige Umfang der Bewertungsmetriken ist begrenzt und reicht nicht aus, um Nachweise für Schlüsselkriterien, wie Verteilungseffekte und dynamische Kosteneffizienz, zu erbringen. Die derzeitige Evaluierungspraxis sollte im Hinblick auf den derzeit begrenzten Umfang der Evaluierungsmetriken weiterentwickelt werden. Wir haben mehrere Aspekte identifiziert, die weiterentwickelt werden könnten, aber am vielleicht dringlichsten erscheint es, stärker auf Verteilungseffekte und die Ausrichtung von Instrumenten zu fokussieren sowie eine dynamische Perspektive bei der Bewertung der Kosteneffizienz einzunehmen. Die Verbesserung dieser beiden Aspekte erfordert nicht nur methodische Verbesserungen, um ein umfassenderes Verständnis der Leistung des Sektors zu gewinnen, sondern auch Reformen zur Verbesserung der Qualität und Zugänglichkeit der Datenerfassung und -bereitstellung. Insbesondere die Auswirkungen auf die Verteilung sollten dringend umfassender berücksichtigt sowie nach Möglichkeit in die Ex-ante-Evaluierung und die Gestaltung der Maßnahmen integriert werden. Werden regressive Ergebnisse nicht ausreichend antizipiert, wird dies wahrscheinlich zu Widerstand gegen die Maßnahmen und zu politischem Druck führen, die Ziele zu lockern. Die derzeitige begrenzte Evaluierung der Auswirkungen des GEG auf einkommensschwache Bevölkerungsgruppen ist zum Beispiel ein wichtiges Versäumnis. Auch die jüngste Evaluierung der BEG zeigt, dass sie vermutlich vorwiegend mittleren und hohen Einkommensgruppen zugutegekommen ist. Eine Neubewertung des Designs und der Ausrichtung der aktuellen Förderprogramme ist erforderlich, während die Umsetzung zusätzlicher Programme,

die speziell auf einkommensschwache Haushalte zugeschnitten sind, um regressive Ergebnisse der Regulierung zu reduzieren (ähnlich wie in Frankreich, Großbritannien und insbesondere dem jüngsten Inflation Reduction Act in den USA), ebenfalls in Betracht gezogen werden könnte.

Bei der derzeitigen Anwendung der Methoden mangelt es in wichtigen Aspekten an Transparenz, was eine eher statische Sichtweise der Emissionsreduzierung und Kosteneffizienz begünstigen kann. Vor allem bei der Anwendung von Emissionsfaktoren mangelt es in den analysierten Berichten an Transparenz. In den veröffentlichten Berichten wird nicht klar erläutert, wie die Emissionsfaktoren zur Vorhersage von Energieeinsparungen verwendet wurden. Unsere Bewertung deutet auf eine statische Anwendung von Emissionsfaktoren hin, welche die sich im Laufe der Zeit entwickelnden Kosteneinsparungen nicht berücksichtigt. Während die Energieeinsparungen über den Lebenszyklus der Maßnahme prognostiziert werden, kann die Verwendung statischer Emissionsfaktoren zu einer unterschätzten Treibhausgasreduktion durch die Installation von Wärmepumpen führen. Dies kann, insbesondere kurzfristig, die energetische Sanierung gegenüber der Dekarbonisierung von Wärmeenergie begünstigen. Eine dynamischere Perspektive sollte über die erwartete THG-Reduktion hinausgehen und Kostenprognosen für die Dekarbonisierung von Strom sowie die erwartete Nachfrage/Kapazität mit potenziellen Einsparungen durch hohe Sanierungsstandards abwägen.

Wir machen mehrere Vorschläge für Reformen, welche die Qualität und den Zugang zu den Daten verbessern würden. Die Einrichtung einer Gebäudedatenbank ist eine vorrangige Reformoption. Die Bundesregierung hat Vorschläge ausgearbeitet, die diese Aufgabe erfüllen würden, und die Renovierungsstrategie als Teil der überarbeiteten EU-Gebäuderichtlinie (sobald sie umgesetzt ist) wird die Einrichtung einer solchen Datenbank erfordern. Die vorrangige Umsetzung von Reformen der Datenmeldung und des Datenschutzes würde die Speicherung aktueller Daten (nicht als Stichproben) ermöglichen und dazu beitragen, die Qualität einer solchen Datenbank in Zukunft zu verbessern. Werden keine sofortigen Maßnahmen ergriffen, besteht die Gefahr, dass es an historischen Daten mangelt, auf die zurückgegriffen werden kann, sobald die Reformen zur Einrichtung einer Datenbank umgesetzt sind.

Bemerkenswert ist der Mangel an Daten über den Energieverbrauch vor der Durchführung der Maßnahmen. Der Datenzugang und die unzuverlässige Aufzeichnungspraxis erschweren eine genaue Bewertung der Auswirkungen der Maßnahmen auf den Energieverbrauch und die Energieeffizienz. Der Energieverbrauch vor der Installation kann überschätzt werden, was dazu führt, dass die Maßnahmen eine höhere Energieeinsparung zu erzielen scheinen als in Wirklichkeit. Dies ist von zentraler Bedeutung für die Berechnung der Energieeinsparungen und der THG-Einsparung sowie der Gesamtwirksamkeit eines Programms. Es bedeutet auch, dass die Bewertung der Kosteneffizienz unzuverlässig sein kann und teure Renovierungsmaßnahmen als praktikabler dargestellt werden können. Darüber hinaus werden aktuelle Anpassungseffekte in der Praxis häufig aufgrund unzureichender Daten ausgeschlossen. Um die Übernahmequoten und die „Zusätzlichkeit“ effektiver zu bewerten, müssen alternative Methoden eingesetzt werden. Derzeit besteht eine Verzerrung durch Selbstauskünfte, da es keine Meldepflicht für nicht-subventionierte Gebäude gibt, was die Erhebung genauer und repräsentativer Daten erschwert.

Vorhersagen und Projektionen hängen in hohem Maße von der Qualität der analytischen Inputs ab. Zwar ist eine gewisse Unsicherheit aufgrund unvorhersehbarer Veränderungen und Schocks unvermeidlich, doch ist die Verbesserung der analytischen Fähigkeit zur Strategieentwicklung von wesentlicher Bedeutung. Dies erfordert die Generierung von Daten, die sowohl vom Umfang her ausreichen, um die potenziell nachteiligsten Auswirkungen der Maßnahmengestaltung zu antizipieren, als auch von strenger methodischer Qualität und gleichzeitig transparent sowie reproduzierbar sind. Angesichts der zunehmenden Bedeutung der Ex-ante-Evaluierung und -Prognose, die durch die jüngsten Änderungen des Bundes-Klimaschutzgesetzes eingeführt wurde, ist die Bewältigung der oben genannten Herausforderungen in Bezug auf Umfang, Datenverfügbarkeit, Durchsetzung und Rechenschaftspflicht von entscheidender Bedeutung, um nützlichere und wirksamere Evaluierungsverfahren zu gewährleisten.

Es ist ein grundlegender Wandel in der Motivation und der wahrgenommenen Rolle der Durchführung von Evaluierungen erforderlich. Anstatt in erster Linie auf die Einhaltung von Berichtspflichten und -anforderungen oder die Rechtfertigung von Bundesausgaben gegenüber dem Bundesrechnungshof abzielen, sollten Evaluierungen von der

Notwendigkeit geleitet sein, zeitnahe und zuverlässige Erkenntnisse für nachfolgende politische Entscheidungen und eine wirksame Neuausrichtung des bestehenden Policy-Mix zu gewinnen. Die Identifizierung und Beseitigung von Engpässen, sei es bei den Verwaltungs-, Umsetzungs-, Durchsetzungs- oder Überwachungskapazitäten oder bei potenziellen Engpässen, die entstehen könnten, kann dazu beitragen, negative unbeabsichtigte Folgen zu mildern. Durch die Verbesserung der Datenqualität, die Einführung umfassender Evaluierungskriterien und die Verstärkung der Durchsetzungsmaßnahmen kann die Steuerung des deutschen Wohnungsbausektors hin zur Klimaneutralität erheblich verbessert werden.

Tabellen

Tabelle 1 - Herausforderungen bei der Gestaltung des Policy-Mix und entsprechende analytische Komponenten für die Bewertung.....	15
Tabelle 3 - In die Inhaltsanalyse einbezogene Dokumente. n = 15.....	28
Tabelle 4 - Interviewteilnehmer.....	29
Tabelle 5 - Formelle Institutionen (Verordnungen) im Wohngebäudebereich, zusammen mit den Anforderungen an die Überwachung, Berichterstattung und Bewertung...	35
Tabelle 6 - KfW-Effizienzstandards für Sanierungen.....	45
Tabelle 7 - Umfang der Ex-post-Evaluierungen.....	64
Tabelle 8 - Verfügbare Fördermittel für KfW-Standards. Quelle: KfW (2023).....	69
Tabelle 9 - Wichtigste Reformoptionen zur Verbesserung des institutionellen Aufbaus.....	110
Tabelle 10 - Wichtigste Reformoptionen zur Verbesserung des Umfangs der bewerteten Programme und der Bewertungsmaßstäbe.....	112
Tabelle 11 - Wichtige Reformoptionen zur Verbesserung des Datenzugangs und der Datenqualität.....	114
Tabelle 12 - Die wichtigsten Reformoptionen zur Verbesserung der Evaluationsmethoden.....	116
Tabelle 13 - Wichtigste Reformoptionen zur Verbesserung der Nutzung von Evaluierungen und der Verbreitung von Informationen.....	117

Abbildungen

Abbildung 1 - Strukturelle Einrichtungen für die Durchführung und Bewertung des BEG und des GEG.....	42
Abbildung 2 - Darstellung von Bottom-up- und Top-down-Ansätzen. Angepasst von Schlomann et al. (2020).....	56
Abbildung 3 - Berichterstattung über das GEG-Programm im zuletzt veröffentlichten NAPE-Bericht.....	58
Abbildung 4 - Metrische Anpassungen, die auf die Endenergieeinsparungen angewendet werden, um die THG-Einsparung in den Methodikleitfäden zu bestimmen.....	65
Abbildung 5 - Berechnung der verursachten privaten Investitionen und der geschaffenen Arbeitsplätze.....	72
Abbildung 6 - Visualisierung der Datengüte/ Belastbarkeit.....	78
Abbildung 7 - Berechnung der Endenergieeinsparungen.....	90
Abbildung 8 - Effektbereinigung zur Berechnung der Nettoeffekte aus dem Brutto.....	92
Abbildung 9 - Verwendung von Brutto-/Nettoeffekten in Methodikleitfäden.....	94
Abbildung 10 - Reformoptionen, geordnet nach Auswirkungen und Vereinbarkeit mit der derzeitigen institutionellen Struktur.....	119

Referenzen

- Altermatt, P.P., Clausen, J., Brendel, H., Breyer, C., Gerhards, C., Kemfert, C., Weber, U., Wright, M., 2023. Der Austausch von Gaskesseln durch Wärmepumpen ist der schnellste Weg, den deutschen Gasverbrauch zu senken. *Commun Earth Environ* 4, 56. <https://doi.org/10.1038/s43247-023-00715-7>
- Auld, G., Mallett, A., Burlica, B., Nolan-Poupart, F., Slater, R., 2014. Evaluating the effects of policy innovations: Lessons from a systematic review of policies promoting low-carbon technology. *Global Environmental Change* 29, 444-458. <https://doi.org/10.1016/j.gloenvcha.2014.03.002>
- Baek, C., Park, S., 2012. Politische Maßnahmen zur Überwindung von Hindernissen bei der energetischen Sanierung bestehender Gebäude. *Renewable and Sustainable Energy Reviews* 16, 3939-3947. <https://doi.org/10.1016/j.rser.2012.03.046>
- BDEW, 2023. Zum Entwurf eines Gesetzes für die Wärmeplanung und zur Dekarbonisierung der Wärmenetze vom 21. Juli 2023. Berlin.
- Berneiser, J., Burkhardt, A., Henger, R., Köhler, B., Meyer, R., Sommer, S., Yilmaz, Y., Kost, C., Herkel, S., 2021. Maßnahmen und Instrumente für eine ambitionierte, klimafreundliche und sozialverträgliche Wärmewende im Gebäudesektor.
- Bovens, M., Schillemans, T., Hart, P., 2008. FUNKTIONIERT DIE ÖFFENTLICHE RECHENSCHAFTSPFLICHT? EIN BEWERTUNGSINSTRUMENT. *Public Adm* 86, 225-242. <https://doi.org/10.1111/j.1467-9299.2008.00716.x>
- Brown, R.R., Farrelly, M.A., Loorbach, D.A., 2013. Actors working the institutions in sustainability transitions: The case of Melbourne's stormwater management. *Global Environmental Change* 23, 701-718. <https://doi.org/10.1016/j.gloenvcha.2013.02.013>
- Bugden, D., Stedman, R., 2019. Ein synthetischer Blick auf die Akzeptanz und das Engagement für intelligente Zähler in den Vereinigten Staaten. *Energy Res Soc Sci* 47, 137-145. <https://doi.org/10.1016/j.erss.2018.08.025>
- Cairney, P., Oliver, K., Wellstead, A., 2016. To Bridge the Divide between Evidence and Policy: Reduce Ambiguity as Much as Uncertainty. *Public Adm Rev* 76, 399-402. <https://doi.org/10.1111/puar.12555>
- Dubash, N.K., Pillai, A.V., Flachsland, C., Harrison, K., Hochstetler, K., Lockwood, M., MacNeil, R., Mildenerger, M., Paterson, M., Teng, F., Tyler, E., 2021. Nationale Klimainstitutionen ergänzen Ziele und Politiken. *Wissenschaft (1979)* 374, 690-693. <https://doi.org/10.1126/science.abm1157>
- Edmondson, D., Flachsland, C., aus dem Moore, N., Koch, N., Koller, F., Brehm, J., Gruhl, H., Levi, S., 2022. Bewertung klimapolitischer Instrumentenmix-Pfade: Eine Anwendung auf den deutschen Leichtfahrzeugsektor. Berlin.
- Edmondson, D., Flachsland, C., Aus Dem Moore, N., Koch, N., Koller, F., Gruhl, H., Brehm, J., 2023. Vorausschauende Klimapolitik-Mix-Pfade: A framework for ex-ante construction and assessment applied to the road transport sector. In Bearbeitung bei Climate Policy.
- Edmondson, D.L., Rogge, K.S., Kern, F., 2020. Zero carbon homes in the UK? Analysis the co-evolution of policy mix and socio-technical system. *Environ Innov Soc Transit* 35, 135-161. <https://doi.org/10.1016/j.eist.2020.02.005>
- ERK, 2022. Zweijahresgutachten 2022 Gutachten zu bisherigen Entwicklungen der Treibhausgasemissionen, Trends der Jahresemissionsmengen und Wirksamkeit von Maßnahmen (gemäß § 12 Abs. 4 Bundes-Klimaschutzgesetz). Berlin.
- ESABCC, 2023. Wissenschaftliche Beratung für die Festlegung eines EU-weiten Klimaziels für 2040 und eines Treibhausgasbudgets für 2030-2050. <https://doi.org/10.2800/609405>
- Finnegan, J.J., 2022. Institutionen, Klimawandel und die Grundlagen langfristiger politischer Entscheidungen. *Comp Polit Stud* 55, 1198-1235. <https://doi.org/10.1177/00104140211047416>
- Fishburn, P.C., 1988. NORMATIVE THEORIES OF DECISION MAKING UNDER RISK AND UNDER UNCERTAINTY, in: *Decision Making*. Cambridge University Press, S. 78-98. <https://doi.org/10.1017/CBO9780511598951.006>

- Fujiwara, N., van Asselt, H., Bößner, S., Voigt, S., Spyridaki, N.-A., Flamos, A., Alberola, E., Williges, K., Türk, A., ten Donkelaar, M., 2019. Die Praxis der Evaluierung von Klimaschutzmaßnahmen in der Europäischen Union und ihren Mitgliedstaaten: Ergebnisse einer Meta-Analyse. *Sustainable Earth* 2, 9. <https://doi.org/10.1186/s42055-019-0015-8>
- Galvin, R., 2023. Politischer Druck zur Nachrüstung von Wohngebäuden in Deutschland mit höheren Energieeffizienzstandards: Ein kosteneffizienter Weg zur Reduzierung der CO₂-Emissionen? *Build Environ* 237, 110316. <https://doi.org/10.1016/j.buildenv.2023.110316>
- Galvin, R., Sunikka-Blank, M., 2016. Quantifizierung von (P)Rebound-Effekten in der Nachrüstungspolitik - Warum ist das wichtig? *Energy* 95, 415-424. <https://doi.org/10.1016/j.energy.2015.12.034>
- Garmston, H., Pan, W., 2013. Non-Compliance with Building Energy Regulations: The Profile, Issues, and Implications on Practice and Policy in England and Wales Non-Compliance with Building Energy Regulations: The Profile, Issues, and Implications on Practice and Policy in England and Wales (Profil, Probleme und Auswirkungen auf Praxis und Politik in England und Wales). *Journal of Sustainable Development of Energy J. sustain. dev. energy water environ. syst* 1, 340-351. <https://doi.org/10.13044/j.sdewes.2013.01.0026>
- George, J.F., Werner, S., Preuß, S., Winkler, J., Held, A., Ragwitz, M., 2023. The landlord-tenant dilemma: Distributional effects of carbon prices, redistribution and building modernisation policies in the German heating transition. *Appl Energy* 339, 120783. <https://doi.org/10.1016/j.apenergy.2023.120783>
- Gillard, R., 2016. Unravelling the United Kingdom's climate policy consensus: The power of ideas, discourse and institutions. *Global Environmental Change* 40, 26-36. <https://doi.org/10.1016/j.gloenvcha.2016.06.012>
- Guy, J., Shears, E., Meckling, J., 2023. Nationale Modelle der Klimagovernance bei den größten Emittenten. *Nat Clim Chang* 13, 748-748. <https://doi.org/10.1038/s41558-023-01688-3>
- Hacker, J., Pierson, P., Thelen, K., 2015. Drift and Conversion: Hidden Faces of Institutional Change. *Advances in Comparative-Historical Analysis*.
- Haug, C., Rayner, T., Jordan, A., Hildingsson, R., Stripple, J., Monni, S., Huitema, D., Massey, E., van Asselt, H., Berkhout, F., 2010. Navigation in den Dilemmata der Klimapolitik in Europa: Erkenntnisse aus Studien zur Bewertung der Politik. *Clim Change* 101, 427-445. <https://doi.org/10.1007/s10584-009-9682-3>
- Hendriks, F., Tops, P., 2003. Reformen der lokalen öffentlichen Verwaltung in den Niederlanden: Modeerscheinungen, Moden und Winde des Wandels. *Public Adm* 81, 301-323. <https://doi.org/10.1111/1467-9299.00348>
- Hildén, M., 2011. Die Entwicklung der Klimapolitik - die Rolle von Lernen und Bewertungen. *J Clean Prod* 19, 1798-1811. <https://doi.org/10.1016/j.jclepro.2011.05.004>
- Hovi, J., Greaker, M., Hagem, C., Holtsmark, B., 2012. Ein glaubwürdiges System zur Durchsetzung der Einhaltung des Klimaregimes. *Climate Policy* 12, 741-754. <https://doi.org/10.1080/14693062.2012.692206>
- Howlett, M., 1998. Vorhersehbare und unvorhersehbare politische Fenster: Institutionelle und exogene Korrelate des kanadischen Agenda-Setting auf Bundesebene. *Canadian Journal of Political Science* 31, 495-524. <https://doi.org/10.1017/S0008423900009100>
- Howlett, M., Ramesh, M., 2003. *Studying Public Policy: Politikzyklen und politische Subsysteme*. Oxford University Press.
- Howlett, M., Ramesh, M., Stewart, J., 1996. Das Studium der öffentlichen Politik: Politikzyklen und politische Subsysteme. *Aust J Polit Sci* 31, 265-266.
- Huitema, D., Jordan, A., Massey, E., Rayner, T., van Asselt, H., Haug, C., Hildingsson, R., Monni, S., Stripple, J., 2011. Die Bewertung der Klimapolitik: Theorie und neue Praxis in Europa. *Policy Sci* 44, 179-198. <https://doi.org/10.1007/s11077-011-9125-7>
- Jacob, K., Kannen, H., 2015. *Klimapolitische Integration im föderalen Rahmen: der Fall der deutschen Baupolitik*. Berlin.

- Jacobs, A.M., Weaver, R.K., 2015. When Policies Undo Themselves: Self-Undermining Feedback as a Source of Policy Change. *Governance* 28, 441-457. <https://doi.org/10.1111/gove.12101>
- Jagnow, K., Wolff, D., 2020. Gutachten: Energetische Gebäudesanierung. Berlin.
- Jordan, A., Huitema, D., Schoenefeld, J., van Asselt, H., Forster, J., 2018. Governing Climate Change Polycentrically, in: *Governing Climate Change*. Cambridge University Press, S. 3-26. <https://doi.org/10.1017/9781108284646.002>
- Kaufmann, W., Hooghiemstra, R., Feeney, M.K., 2018. Formelle Institutionen, informelle Institutionen und Bürokratie: A comparative study. *Public Adm* 96, 386-403. <https://doi.org/10.1111/padm.12397>
- Kern, F., 2011. Ideen, Institutionen und Interessen: Zur Erklärung politischer Divergenzen bei der Förderung von "Systeminnovationen" in Richtung Nachhaltigkeit. *Environ Plann C Gov Policy* 29, 1116-1134. <https://doi.org/10.1068/c1142>
- Köhler, J., Geels, F.W., Kern, F., Markard, J., Onsongo, E., Wieczorek, A., Alkemade, F., Avelino, F., Bergek, A., Boons, F., Fünfschilling, L., Hess, D., Holtz, G., Hyysalo, S., Jenkins, K., Kivimaa, P., Martiskainen, M., Mcmeekin, A., Susan, M., Nykvist, B., Pel, B., Raven, R., Rohrer, H., Sandén, B., Schot, J., Sovacool, B., Turnheim, B., Welch, D., Wells, P., 2019. Eine Agenda für die Forschung zu Nachhaltigkeitsübergängen: State of the art and future directions ☆. *Environ Innov Soc Transit* 1-32. <https://doi.org/10.1016/j.eist.2019.01.004>
- Krippendorff, K., 2004. Verlässlichkeit in der Inhaltsanalyse. *Hum Commun Res* 30, 411-433. <https://doi.org/10.1111/j.1468-2958.2004.tb00738.x>
- Kuzemko, C., 2016. Entpolitisierung der Energiepolitik im Vereinigten Königreich: Entpolitisierung Zerstörung der politischen Energie im Vereinigten Königreich: <https://doi.org/10.1111/1467-856X.12068>
- Levi, S., 2021. Warum hassen wir Kohlenstoffsteuern? Erkenntnisse des maschinellen Lernens über die Rolle der persönlichen Verantwortung, des Vertrauens, der Rückführung von Einnahmen und anderer Faktoren in 23 europäischen Ländern. *Energy Res Soc Sci* 73, 101883. <https://doi.org/10.1016/j.erss.2020.101883>
- Lockwood, M., Kuzemko, C., Mitchell, C., Hoggett, R., 2017. Historischer Institutionalismus und die Politik der nachhaltigen Energiewende: A research agenda. *Environ Plann C Gov Policy* 35, 312-333. <https://doi.org/10.1177/0263774X16660561>
- Lu, Y., Karunasena, G., Liu, C., 2022. A Systematic Literature Review of Non-Compliance with Low-Carbon Building Regulations. *Energies (Basel)* 15, 9266. <https://doi.org/10.3390/en15249266>
- Magro, E., Wilson, J.R., 2017. Governing the Evaluation of Policy Mixes in the Context of Smart Specialisation Strategies: Emerging Challenges First Draft.
- Mastenbroek, E., van Voorst, S., Meuwese, A., 2016. Closing the regulatory cycle? Eine Meta-Evaluierung von Ex-post-Evaluierungen von Rechtsvorschriften durch die Europäische Kommission. *J Eur Public Policy* 23, 1329-1348. <https://doi.org/10.1080/13501763.2015.1076874>
- Mayring, P., 2000. Qualitative Inhaltsanalyse.
- Meckling, J., Nahm, J., 2018. The power of process: State capacity and climate policy. *Governance* 31, 741-757. <https://doi.org/10.1111/gove.12338>
- Melvin, J., 2018. The split incentives energy efficiency problem: Evidence of underinvestment by landlords. *Energy Policy* 115, 342-352. <https://doi.org/10.1016/j.enpol.2017.11.069>
- Meyer, R., Berneiser, J., Burkhardt, A., Doderer, H., Eickelmann, E., Henger, R., Köhler, B., Sommer, S., Yilmaz, Y., Blesl, M., Bürger, V., Braungardt, S., 2021. Maßnahmen und Instrumente für eine ambitionierte, klimafreundliche und sozialverträgliche Wärmewende im Gebäudesektor: Teil 2: Instrumentensteckbriefe für den Gebäudesektor.
- Michaelowa, A., 2008. Deutsche Klimapolitik zwischen Global Leadership und Muddling Through, in: *Turning Down the Heat*. Palgrave Macmillan UK, London, S. 144-163. https://doi.org/10.1057/9780230594678_9

- Michaelowa, A., Allen, M., Sha, F., 2018. Politische Instrumente zur Begrenzung des globalen Temperaturanstiegs auf 1,5°C - kann die Menschheit die Herausforderung meistern? *Climate Policy* 18, 275-286. <https://doi.org/10.1080/14693062.2018.1426977>
- Moore, T., Doyon, A., 2023. Facilitating the Sustainable Housing Transition, in: *A Transition to Sustainable Housing*. Springer Nature Singapur, Singapur, S. 239-258. https://doi.org/10.1007/978-981-99-2760-9_8
- Morrison, T.H., 2022. Die Gestaltung vorausschauenderer politischer Pfade. *Nat Sustain* 5, 372-373. <https://doi.org/10.1038/s41893-022-00894-9>
- Mukherjee, I., Coban, M.K., Bali, A.S., 2021. Politische Kapazitäten und effektive Maßnahmengestaltung: ein Überblick. *Policy Sci* 54, 243-268. <https://doi.org/10.1007/s11077-021-09420-8>
- Öberg, P., Lundin, M., Thelander, J., 2015. Politische Macht und Maßnahmengestaltung: Why Are Policy Alternatives Constrained? *Policy Studies Journal* 43, 93-114. <https://doi.org/10.1111/psj.12086>
- Patashnik, E.M., 2009. *Reforms at Risk*. Princeton University Press. <https://doi.org/10.1515/9781400828852>
- Peters, B.G., 2012. *Institutionelle Theorie in der Politikwissenschaft: der neue Institutionalismus / B. Guy Peters*. Continuum, New York.
- Qian, Q.K., Fan, K., Chan, E.H.W., 2016. Regulatorische Anreize für grüne Gebäude: Konzessionen für die Bruttogeschossfläche. *Building Research & Information* 44, 675-693. <https://doi.org/10.1080/09613218.2016.1181874>
- Rosenow, J., Galvin, R., 2013. Evaluating the evaluations: Evidence from energy efficiency programmes in Germany and the UK. *Energy Build* 62, 450-458. <https://doi.org/10.1016/j.enbuild.2013.03.021>
- Rosenow, J., Kern, F., Rogge, K., 2017. Die Notwendigkeit eines umfassenden und gezielten Instrumentenmixes zur Förderung der Energiewende: The case of energy efficiency policy. *Energy Res Soc Sci* 33. <https://doi.org/10.1016/j.erss.2017.09.013>
- Sabatier, P., Weible, C., 2014. *Theories of the Policy Process*. Routledge.
- Schmidt, V.A., 2008. Diskursiver Institutionalismus: The Explanatory Power of Ideas and Discourse. *Annual Review of Political Science* 11, 303-326. <https://doi.org/10.1146/annurev.polisci.11.060606.135342>
- Schoenefeld, J., Jordan, A., 2017. Governing policy evaluation? Towards a new typology. *Evaluation* 23, 274-293. <https://doi.org/10.1177/1356389017715366>
- Singhal, P., Pahle, M., Kalkuhl, M., Levesque, A., Sommer, S., Berneiser, J., 2022. Beyond good faith: Warum eine evidenzbasierte Politik notwendig ist, um Gebäude in Deutschland kosteneffizient zu dekarbonisieren. *Energy Policy* 169, 113191. <https://doi.org/10.1016/j.enpol.2022.113191>
- Skogstad, G., 2023. Historischer Institutionalismus in der öffentlichen Politik, in: *Encyclopedia of Public Policy*. Springer International Publishing, Cham, S. 1-9. https://doi.org/10.1007/978-3-030-90434-0_21-1
- Steinmo, S., Thelen, K., 1992. *Die Strukturierung der Politik: Historischer Institutionalismus in der vergleichenden Analyse*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511528125>
- Sunikka-Blank, M., Galvin, R., 2012. Einführung des Prebound-Effekts: die Lücke zwischen Leistung und tatsächlichem Energieverbrauch. *Building Research & Information* 40, 260-273. <https://doi.org/10.1080/09613218.2012.690952>
- Thelen, K., 1999. HISTORISCHER INSTITUTIONALISMUS IN DER VERGLEICHENDEN POLITIKWISSENSCHAFT 369-404.
- Tosun, J., De Francesco, F., Peters, B.G., 2019. Von umweltpolitischen Konzepten zu praktikablen Instrumenten: Wissensbildung und -delegation in Mehrebenensystemen. *Public Adm* 97, 399-412. <https://doi.org/10.1111/padm.12544>
- Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science* (1979) 211, 453-458. <https://doi.org/10.1126/science.7455683>

- van den Bergh, J., Castro, J., Drews, S., Exadaktylos, F., Foramitti, J., Klein, F., Konc, T., Savin, I., 2021. Gestaltung eines effektiven klimapolitischen Mixes: Berücksichtigung von Synergieeffekten der Instrumente. *Climate Policy*. <https://doi.org/10.1080/14693062.2021.1907276>
- von Lüpke, H., Leopold, L., Tosun, J., 2022. Institutionelle Koordinationsarrangements als Elemente von Maßnahmengestaltungsräumen: Erkenntnisse aus der Klimapolitik. *Policy Sci*. <https://doi.org/10.1007/s11077-022-09484-0>
- Wang, Q., Hubacek, K., Feng, K., Wei, Y.-M., Liang, Q.-M., 2016. Distributional effects of carbon taxation. *Appl Energy* 184, 1123-1131. <https://doi.org/10.1016/j.apenergy.2016.06.083>
- Weber, K.M., Rohracher, H., 2012. Legitimierung von Forschungs-, Technologie- und Innovationspolitik für transformativen Wandel: Die Kombination von Erkenntnissen aus Innovationssystemen und Mehrebenenperspektive in einem umfassenden "Versagensrahmen". *Res Policy* 41, 1037-1047. <https://doi.org/10.1016/j.respol.2011.10.015>
- Weiß, J., Vogelpohl, T., 2010. Politische Instrumente zur Erhöhung der energetischen Sanierungsquote bei Eigenheimen.
- Wurzel, R.K.W., 2010. Umwelt-, Klima- und Energiepolitik: Pfadabhängiger Inkrementalismus oder Quantensprung? *Ger Polit* 19, 460-478. <https://doi.org/10.1080/09644008.2010.515838>
- Zachmann, G., Fredriksson, G., Claeys, G., 2018. DIE VERTEILUNGSEFFEKTE DER KLIMAPOLITIK.

Anhang I: Geprüfte Dokumente

Eingeschlossene Typen		Titel	Jahr	Autoren	Kunde
Nein	Ex-ante-Evaluierung	Kurzgutachten zur Überarbeitung von Anforderungssystemen und Standards im Gebäudeenergiegesetz für Neubauten sowie Bestandsgebäude einschl. der Wirtschaftlichkeitsbetrachtungen für Neubauten und Bestandsgebäude	2022	Dena, Guidehouse, iTG, Öko-Institut, Prognos, FIW München, ifeu	BMWK
Nein	Ex-ante-Evaluierung	Einsparpotenziale aus der Optimierung von Heizungsanlagen in Wohngebäuden	2022	Borderstep Institut, Dena	BMWK
Nein	Ex-ante-Evaluierung	Mindestvorgaben für die Gesamteffizienz von Bestandsgebäuden	2022	bbh, ifeu, Öko-Institut, Prognos	BfEE
Ja	Ex-post-Evaluierung	Evaluation und Perspektiven des Marktanzreizprogramms zur Förderung von Maßnahmen zur Nutzung erneuerbarer Energien im Wärme- markt im Förderzeitraum 2019 bis 2020	n.d.	ifeu, Fraunhofer, Fichtner GmbH	BMWK
Ja	Ex-post-Evaluierung	Abschlussbericht zur Evaluation der Richtlinie über die Förderung der Heizungsoptimierung durch hocheffiziente Pumpen und hydraulischen Abgleich	2022	Arepo, Wuppertal Institut	BfEE
Ja	Ex-post-Evaluierung	Bewertung des Förderprogramms KfW 433	2022	Prognos	BMWK
Ja	Ex-post-Evaluierung	Bewertung der Förderprogramme der EBS WG im Förderzeitraum 2018	2022	Prognos, FIW München	BMWK
Ja	Ex-post-Evaluierung	Förderwirkungen BEG EM 2021	2023	Prognos, ifeu, FIW München, iTG	BMWK
Ja	Ex-post-Evaluierung	Förderwirkungen BEG WG 2021	2023	Prognos, ifeu, FIW München, iTG	BMWK
Nein	Ex-post-Evaluierung	Überwachung der Initiative Energieeffizienz-Netzwerke	2022	Adelphi, Fraunhofer	BMWK
Ja	Leitfäden	Methodikpapier zur ex-ante Abschätzung der Energie- und THG-Einsparungswirkung von energie- und klimaschutzpolitischen Maßnahmen	2022	Prognos, Fraunhofer ISI, Öko-Institut e.V.	BMWK
Ja	Leitfäden	Methodikleitfaden für Evaluationen von Energie-Effizienzmaßnahmen des BMWi (Projekt Nr. 63/15 - Aufstockung)	2020	Prognos, ifeu, Stiftung Umweltenergierecht, Fraunhofer ISI	BMWK
Nein	Nationales Monitoring	DENA-GEBÄUDEREPORT 2023	2023	Dena	BMWK
Ja	Nationales Monitoring	Klimaschutz in Zahlen	2022	BMWK	n.a.

Ja	Nationales Monitoring	Klimaschutzbericht 2022	2022	BMWK	n.a.
Ja	Nationales Monitoring	Zweijahresgutachten 2022 Gutachten zu bisherigen Entwicklungen der Treibhausgasemissionen, Trends der Jahresemissionsmengen und Wirksamkeit von Maßnahmen	2022	Expertenrat für Klimafragen	n.a.
Ja	Nationales Monitoring	Die Energie der Zukunft 8 th Monitoringbericht zur Energiewende - Berichtsjahre 2018 und 2019	2021	BMWK	n.a.
Ja	Nationales Monitoring	Energieeffizienz in Zahlen	2021	BMWK	n.a.
Ja	Andere	Begleitung von BMWK-Maßnahmen zur Umsetzung einer Wärmepumpen-Offensive	2023	Dena, Guidehouse, iTG, Öko-Institut, Prognos, EY, pwc, bbh, FIW München, ifeu, heimrich + hannot	BMWK
Nein	Andere	Metastudie zur Verbesserung der Datengrundlage im Gebäudebereich	2022	bbh, dena, EY, FIW München, heimrich+hannot	BMWK
Nein	Vierteljährliche Berichterstattung	Bundesförderung für effiziente Gebäude (BEG) Berichterstattung zur BEG-Förderung im 4. Quartal 2022 (Stand: 31.12.2022)	2022	BMWK	n.a.
Nein	Strategie	Sofortprogramm gemäß § 8 Abs. 1 KSG für den Sektor Gebäude	2022	BMWK, BMWSB	n.a.
Nein	Strategie	Hintergrundpapier zur Gebäudestrategie Klimaneutralität 2045	2022	Dena, Guidehouse, iTG, Öko-Institut, Prognos, EY, adelphi, bbh, FIW München, ifeu	BMWK
Nein	Strategie	Klimaneutrales Deutschland 2045	2021	Prognos, Öko-Institut, Wuppertal-Institut	Stiftung Klimaneutralität, Agora Energiewende, Agora Verkehrs-wende
Nein	Strategie	Energieeffizienzstrategie 2050	2019	BMWK	n.a.
Ja	Strategie	Nationaler Aktionsplan für Energieeffizienz	2014	BMWK	n.a.
Nein	Strategie	Klimapfade 2.0 - Ein Wirtschaftsprogramm für Klima und Zukunft	2021	BCG	BDI
Nein	Strategie	Aufbruch Klimaneutralität	2021	EWI, FIW, ITG, (...)	Dena
Nein	Strategie	Langfristszenarien für die Transformation des Energiesystems in Deutschland	2022	Consentec, TU Berlin, Ifeu	BMWK

Nein	Strategie	Deutschland auf dem Weg zur Klimaneutralität 2045	2021	PIK, MCC, PSI, RWI, IER, (...)	Ariadne - Koper-nikus-Projekte
Nein	Jährlicher Sub-ven-tionsber-icht	Förderreport KfW Bankengruppe	2022	KfW	n.a.

Tabelle 14: Geprüfte Dokumente, n = 30

Anhang II: Fragebogen für halbstrukturierte Interviews

A. Prozess und Koordination

1. Welche Ex-post-Evaluierungsstudien werden von den Ministerien in Auftrag gegeben?
2. Welche Agenturen und Beratungsunternehmen beauftragen sie?
3. Inwieweit koordinieren die Ministerien die Beauftragung und den Einsatz dieser Instrumente?
4. Werden nicht in Auftrag gegebene Studien berücksichtigt?

B. Qualität und Umfang

5. Werden einzelne Instrumente oder Pakete bewertet? Werden bei der Evaluierung von Paketen auch die Wechselwirkungen zwischen den Instrumenten analysiert?
6. Sind Sie der Meinung, dass sie im Hinblick auf die Abdeckung wichtiger Bewertungskriterien wie Treibhausgasreduzierung, Verbesserung der Energieeffizienz, Kosten, Nutzen und Verteilungseffekte angemessen sind?
7. Sind Sie der Meinung, dass die Ex-post-Evaluierung der Politik von angemessener Qualität ist?
8. Welche Art von Ex-post-Evaluierungsstudien würden Sie für nützlich halten - in Bezug auf die berücksichtigten Evaluierungskriterien, den Umfang der Gebäude und Technologien oder andere relevante Aspekte?

C. Auswirkungen auf den politischen Prozess (Reformen)

9. Welche Auswirkungen haben Ihrer Erfahrung nach Ex-post-Evaluierungsstudien auf die Maßnahmengestaltung und Reformen?
10. Allgemeiner ausgedrückt: Auf welchen analytischen Input stützen sich die Reformen (Reformvorschläge)? Intern oder von Beratungsunternehmen?
11. Wenn Sie einen Vorschlag machen könnten, um den Rahmen für Ex-post-Evaluierungsstudien von Grund auf (oder sogar schrittweise) neu zu gestalten, wie würden Sie dabei vorgehen?

Anhang III: Zusätzliche Informationen zu Gesetzgebungen

Klimaaktionsprogramm 2020

Das Klimaaktionsprogramm 2020 wurde erstmals 2014 vorgestellt. Es enthielt 110 Maßnahmen. Für jede Maßnahme wurde die Menge an THG-Emissionen ermittelt, die bis 2020 eingespart werden sollte (BRH 2022: 16). Der Umsetzungsstand und die erwarteten Wirkungen der Maßnahmen wurden jährlich überprüft (sogenannte *Quantifizierungsberichte*). Die Ergebnisse wurden im Klimaschutzbericht der Bundesregierung veröffentlicht (BRH 2022: 16).

Die Bewertung durch den Bundesrechnungshof (BRH) ergab, dass viele Maßnahmen des Aktionsprogramms Klimaschutz 2020 wirkungslos blieben (BRH 2022: 7). Die Bundesregierung hat sich die Ergebnisse der Quantifizierungsberichte und andere Prognosen zur THG-Minderung wiederholt "*nicht zu eigen gemacht*" (BRH 2022: 16). Der BRH argumentierte, dass die Berechnungen nur eine grobe Abschätzung der Wirkung mit Blick auf das Zieljahr (ex-ante) darstellen und eine detaillierte und empirisch gestützte Bewertung der einzelnen Maßnahmen (ex-post) nicht ersetzen würden (BRH 2022: 16). Laut BRH haben nur 8 von 110 Maßnahmen signifikant zur THG-Minderung beigetragen, während 70 Maßnahmen keinen Beitrag zur THG-Minderung leisteten. Sie werden als Begleitinstrumente betrachtet (BRH 2022: 16). Diese Einschätzung erklärt sich zum Teil aus der Kritik des BRH, dass für viele Instrumente keine quantifizierbaren THG-Einsparungen vorliegen.

Sofortprogramm für den Klimaschutz 2022

Das Klimaschutz-Sofortprogramm 2022 enthält kaum neue Maßnahmen, stattdessen werden die Mittel für bestehende Programme erhöht. Die Bundesregierung beabsichtigt, die Finanzierung wichtiger klimapolitischer Projekte zwischen 2022 und 2025 mit mehr als 93 Milliarden Euro (BReg 2021) deutlich auszuweiten. Mehr als die Hälfte der zusätzlichen Mittel im Sofortprogramm ist für die Förderung der energetischen Gebäudesanierung und des Einbaus energieeffizienter Heizungsanlagen vorgesehen (BReg 2021). Allein 4,5 Milliarden Euro sollen zusätzlich zur Verfügung gestellt werden (BReg 2021). Ab 2023 will die Bundesregierung keine Heizungsanlagen mehr fördern, die ausschließlich mit fossilen Energieträgern betrieben werden (BReg 2021). Einschätzung des BRH, dass das Klimaschutz-Sofortprogramm 2022 nur wenige neue Maßnahmen enthält und weitgehend aus Mittelaufstockungen für bestehende Programme oder bloßen Absichtserklärungen besteht (BRH 2022: 20).

Das BMU hält eine Ex-ante-Evaluierung der THG-Einsparungen einzelner Maßnahmen für geeignet, um die Wirkung von Klimaschutzmaßnahmen zuverlässig und vergleichbar zu bewerten. Allerdings ist eine solche Bewertung nicht immer möglich (BRH 2022: 22). Ein kontinuierliches Monitoring der THG-Minderung einzelner Maßnahmen über die jährliche Klimaschutzberichterstattung hinaus ist nicht zielführend. Das BMF teilt die Position des BMU und weist zusätzlich darauf hin, dass die Ausgestaltung und Umsetzung des Sofortprogramms Klimaschutz 2022 im Hinblick auf die Nähe zur Bundestagswahl unterlassen wurde (BRH 2022: 22).

Nationaler Aktionsplan für Energieeffizienz

Der NAPE (Nationaler Aktionsplan Energieeffizienz) wurde als umfassendes Maßnahmenpaket zur Verbesserung der Energieeffizienz in Deutschland konzipiert. Der NAPE wurde erstmals 2014 zusammen mit dem Klimaaktionsplan 2020 umgesetzt. Er gibt die strategische Ausrichtung der Energieeffizienzpolitik mit drei zentralen Zielen vor:

- Förderung der Energieeffizienz im Gebäudesektor: Der ursprüngliche NAPE betonte die Bedeutung der Verbesserung der Energieeffizienz in Gebäuden als Schlüsselbereich für die Erzielung von Energieeinsparungen und die Reduzierung von Emissionen.
- Etablierung von Energieeffizienz als Ertrags- und Geschäftsmodell: Der NAPE zielte darauf ab, Energieeffizienz als rentable Geschäftsmöglichkeit zu fördern und Investitionen in energieeffiziente Technologien und Praktiken anzuregen.
- Stärkung der Eigenverantwortung für Energieeffizienz: Der NAPE sollte das Gefühl der individuellen und kollektiven Verantwortung für die Energieeffizienz fördern und Bürger, Unternehmen und Organisationen ermutigen, proaktive Maßnahmen zur Reduzierung des Energieverbrauchs zu ergreifen.

Der NAPE 2.0 wurde im Jahr 2019 als Teil der Energieeffizienzstrategie 2050 (BMWi 2019) umgesetzt. Er berücksichtigt politische Erkenntnisse und passt sich an neue Entwicklungen an, insbesondere mit Blick auf den Zeitraum von 2021 bis 2030. Zu den wichtigsten Unterschieden zwischen dem NAPE 2.0 und dem ursprünglichen NAPE gehören:

- Stärkere Konzentration auf die Nachfrageseite des Energiesystems: Der NAPE 2.0 verlagert seinen Schwerpunkt vom Gebäudesektor auf eine breitere Perspektive der Energieeffizienz, wobei der Schwerpunkt auf der Nachfrageseite und auf Maßnahmen zur Reduzierung des Energieverbrauchs in der gesamten Wirtschaft liegt.

- Integration von Energieeffizienzmaßnahmen aus dem Klimaschutzprogramm 2030: Der NAPE 2.0 bezieht relevante Maßnahmen und Ziele aus dem Klimaschutzprogramm 2030 ein, um die Ziele der Energieeffizienz mit den umfassenderen Klimaschutzzielen in Einklang zu bringen.
- Spezifischer Fokus auf die Reduzierung des Endenergieverbrauchs zwischen 2021 und 2030: Der NAPE 2.0 legt explizite Ziele und Maßnahmen fest, um die Reduzierung des Energieverbrauchs während des angegebenen Zeitraums voranzutreiben und zu einer allgemeinen Verbesserung der Energieeffizienz beizutragen.
- Einführung eines jährlichen Überwachungsprozesses: Der NAPE 2.0 führt einen Überwachungsmechanismus ein, um den Erfolg und die Wirksamkeit seiner Maßnahmen regelmäßig zu bewerten. Diese fortlaufende Überwachung ermöglicht es, notwendige Anpassungen oder Änderungen zu ermitteln, um die Ergebnisse des Plans zu verbessern.

Diese Anpassungen und Verbesserungen im NAPE 2.0 spiegeln das sich entwickelnde Verständnis für die Herausforderungen und Chancen der Energieeffizienz sowie die Notwendigkeit wider, sich an langfristigen Klima- und Nachhaltigkeitszielen zu orientieren. Der NAPE 2.0 wird in diesem Jahr erneut aktualisiert und auf die Energieeffizienz für eine klimaneutrale Zukunft 2045 abgestimmt.

Die Daten, die in die NAPE-Berichterstattung einfließen, konzentrieren sich auf Energieeinsparungen und die Verringerung von Treibhausgasen. Mindestens die folgenden Daten müssen erhoben werden, um die Vorlage für die NAPE-Überwachung auszufüllen (Schlomann et al. 2020: 50-51):

- Jährlich werden neue Werte für die folgenden Kategorien hinzugefügt:
 - Endenergieeinsparungen bei Strom
 - Endenergieeinsparungen bei Kraftstoffen
 - Alternativ: aggregierte Werte, aber individuelle Darstellung bevorzugt
 - Lebenslang
 - Subventionsvolumen
 - Ausgelöste Investitionen
 - Anzahl der Fälle von Beförderung oder sonstiger Tätigkeitsparameter
- Wie die Einsparungen ermittelt wurden
- Erläuterung der Wahl der Basislinie
- Anpassungsfaktoren für die Effektbereinigung

Diese Berichterstattung erstreckt sich nicht auf andere wichtige Kriterien wie makroökonomische Auswirkungen, Verteilungseffekte/sozioökonomische Auswirkungen oder administrative Herausforderungen oder Kosten (Ex-post-Leitfaden, S. 50-51).

Bundesförderung für energieeffiziente Gebäude (BEG)

Die Förderung der Sanierung von Wohngebäuden umfasst die Gebäudehülle, die Anlagentechnik, die Wärmeerzeugungsanlagen, die Heizungsoptimierung sowie die technische Planung und Überwachung. Die BEG umfasst auch andere Maßnahmen (z.B. Neubau), die jedoch nicht direkt auf die Sanierung von Wohngebäuden zutreffen und daher in diesem Bericht nicht behandelt werden.

Einzelmaßnahmen an der Gebäudehülle tragen zur Steigerung der Energieeffizienz des Gebäudes bei, wie z.B. Fenster oder Türen und Dämmung der Außenwände oder des Daches. *Anlagentechnik* (außer Heizungstechnik), umfasst die Installation von Anlagentechnik in bestehenden Gebäuden zur Steigerung der Energieeffizienz des Gebäudes, wie z.B. eine energieeffiziente Lüftungs- und Klimatisierungsanlage. *Wärmeerzeugungsanlagen* (Heizungstechnik) fördert den Einbau von effizienten Wärmeerzeugern, Anlagen zur Heizungsunterstützung und den Anschluss an ein Gebäude- oder Wärmenetz, das erneuerbare Energien zur Wärmeerzeugung mit einem Anteil von mindestens 25 Prozent einbindet. Die *Heizungsoptimierung* fördert Maßnahmen zur Optimierung des Wärmeverteilsystems einer Wärmeerzeugungsanlage in bestehenden Gebäuden (mindestens 2 Jahre alt) und erhöht die Energieeffizienz der Anlage, wie z.B. den hydraulischen Abgleich oder den Austausch der Heizungspumpe. Die *Fachplanung und Baubegleitung* erbringt energetische Fachplanungs- und Baubegleitungsleistungen im Zusammenhang mit der Umsetzung von geförderten Maßnahmen im Sinne dieses Förderprogramms.

Gebäudeenergiegesetz (GEG)

Nachrüstpflichten. Das GEG sieht für bestimmte Teile (Austausch bestimmter alter Heizkessel, Dämmung bestimmter Rohrleitungen, Dämmung der obersten Geschossdecke, Einbau bestimmter Regelungstechnik von Heizungs- und Klimaanlage) sogenannte Nachrüstpflichten unabhängig von Maßnahmen vor. Weitere Nachrüstpflichten sind Gegenstand von §47 (oberste Geschossdecke), §61(1) (zentrale Regeleinrichtungen für zentrale Heizungsanlagen), §63(3) (raumweise Regeleinrichtungen für zentrale Heizungsanlagen), §66 (Regeleinrichtungen für Raumluftfeuchte) und §§71 bis 73 GEG (Rohrleitungsdämmung, Außerbetriebnahme alter Heizkessel). Nach §46 GEG dürfen Außenbauteile eines Gebäudes nicht so verändert werden, dass die energetische Qualität eines Gebäudes verschlechtert wird. Eine Ausnahme gilt für Änderungen,

wenn die Fläche der geänderten Bauteile nicht mehr als 10 % der Gesamtfläche der Bauteilgruppe nach Anlage 7 betrifft. Darüber hinaus gelten die Anforderungen für bestehende Gebäude nicht, wenn die Einhaltung mit anderen Vorschriften über Gesundheit, Lärm, Brandschutz, Arbeitssicherheit und Stabilität kollidiert.

Bedingte Anforderungen für die Sanierung (§ 48 GEG). Nach § 48 GEG werden an Umbauten nach Anlage 7 GEG Anforderungen an die energetische Qualität gestellt. Sind an einem Bauteil ohnehin größere Maßnahmen erforderlich (z.B. aus Gründen der Bauunterhaltung oder der Sorgfaltspflicht) oder aus anderen Gründen geplant, müssen auch für die betroffenen Teilbereiche energetische Standards eingehalten werden. Die Bauteilanforderungen nach § 48 GEG in Verbindung mit Anlage 7 GEG müssen eingehalten werden, wenn bestimmte Maßnahmen durchgeführt werden. Die entsprechenden Maßnahmen, die mit Auflagen verbunden sind, sind in der zweiten Spalte der Tabelle in Anlage 7 GEG beschrieben. Bei der Erneuerung von Außenwänden, Bauteilen im Dachbereich, Wänden gegen unbeheizte Räume oder das Erdreich und Deckenflächen, die nach unten von unbeheizten Räumen, der Außenluft oder dem Erdreich getrennt sind, gelten die bedingten Anforderungen nicht für Bauteile, die nach dem 31. Dezember 1983 (d.h. nach Inkrafttreten der Wärmeschutzverordnung 2nd) unter Einhaltung der Energieeinsparverordnung errichtet oder erneuert wurden.

Der Gesamtnachweis kann auf das Gebäude angewendet werden. Anstelle der Einhaltung der Bauteilanforderungen für alle relevanten Änderungen an der Gebäudehülle kann auch ein Gesamtnachweis nach § 50 für das Gebäude geführt werden. Im Gesamtnachweis nach § 50 ist nachzuweisen, dass ein Wohngebäude nach der Sanierung den Primärenergiebedarf des Referenzgebäudes nach Anlage 1 GEG um nicht mehr als 40 % überschreitet und dass sein spezifischer Transmissionswärmebedarf nach § 50 Absatz 2 begrenzt ist. Wird bei der Sanierung eines Gebäudes gleichzeitig eine Erweiterung des bestehenden Gebäudeteils durchgeführt und wird der Gesamtnachweis nach § 50 GEG gewählt, sind die Berechnungen für das gesamte Gebäude durchzuführen. Auf der Grundlage von § 79 (2) GEG ist in diesem Fall ein Energieausweis für das gesamte Gebäude auszustellen und auf Seite 1 als Anlassmodernisierung zu kennzeichnen. Die Einhaltung der Anforderungen für den neuen Gebäudeteil ist eigenständig und gesondert nachzuweisen.

Anhang IV: Empfohlene Verfahren

Box 1: Ex-post-Evaluierungen - empfohlenes Verfahren


Schritt:

1. Charakterisierung des Förderprogramms
 - Allgemeiner Überblick über die Maßnahme, einschließlich Zielgruppe/Sektoren, Budget, Finanzierungs-/Durchführungsstellen, Rechtsgrundlage, damit zusammenhängende Politiken und Finanzierungsverfahren
 - Entwicklung und Analyse des Wirkungsmodells der Maßnahme
2. Annahmen zu den wichtigsten Rahmendaten für die Folgenabschätzung
 - Bestimmung der Rahmenannahmen der bereitgestellten Daten, die für die Folgenabschätzung relevant sind
3. Identifizieren der Ziele des Programms
 - Beschreibung der Anforderungen und Erwartungen an die Evaluation
 - Analyse der Top-Down-Ziele für das Förderprogramm anhand von öffentlichen Dokumenten, Richtlinien und Gesetzen
 - Analyse der Bottom-up-Ziele des Förderprogramms auf der Grundlage von Ex-ante-Schätzungen und Förderrichtlinien
 - Definition der wichtigsten Bereiche, die für die Bewertung von Interesse sind
4. Definition von Indikatoren zur Messung der Zielerreichung
 - Auswahl von Indikatoren, die den Fortschritt in den relevanten Bereichen der Evaluierung widerspiegeln (Zielerreichung, Wirkungs- und Effizienzkontrolle)
 - Operationalisierung der Indikatoren: Wahl zwischen qualitativem/quantitativem Typ, Beschreibung und Abgrenzung, Berechnungsmodell, Art des Ergebnisses, Einheiten (quantitativ) oder Skalen einschließlich Interpretationsregeln (qualitativ).
5. Datenerhebung
 - Erstellung eines Erhebungskonzepts auf Basis der Auswahl und Struktur der Indikatoren
 - Durchführung der Datenerhebung
6. Methodisches Verfahren zur Ermittlung der (Brutto-)Auswirkungen der Maßnahme
 - Festlegung der Berechnungsmethode zur Ermittlung der Bruttoauswirkungen
 - Berechnung des Bruttoeffekts
7. Methodisches Verfahren zur Anpassung der Wirkung (Nettoeffekt der Maßnahme)
 - Festlegung der Methodik für die Effektbereinigung
 - Durchführung der Effektbereinigung: Bereinigung des Bruttoeffekts um Mitnahme-, Vorzieh-, Spillover-, Verzögerungs-, Struktur- und Rebound-Effekte.

Box 2: Empfohlene Schritte der Ex-ante-Evaluierung:

Schritt:

1. Berücksichtigte Bewertungskriterien und anzuzeigende Indikatoren
 - Festlegung der zu berücksichtigenden Bewertungskriterien: neben dem Energie- und THG-Einsparungseffekt je nach Bewertungsziel Berücksichtigung weiterer Kriterien wie ökonomische Effekte (Auswirkungen auf Energiekosten, Investitionen, Beschäftigung, Wertschöpfung), soziale Effekte oder Akzeptanzfragen.
 - Definition von quantitativen oder qualitativen Indikatoren für die konkrete Bewertung einer Maßnahme: Für das Kriterium der Einspareffekte sind dies die End-, Primär- und THG-Einsparungen, wobei unterschiedliche Berechnungsmodi (neue / zusätzliche jährliche Einsparungen, zusätzliche Einsparungen über einen Zeitraum) zu berücksichtigen sind.
2. Annahmen zu den wichtigsten Rahmendaten für die Folgenabschätzung
 - Festlegung der für die Folgenabschätzung relevanten Rahmenannahmen zu Energiepreisen, Lebensdauern sowie Emissions- und Primärenergiefaktoren.
 - Festlegung einer Referenz (Baseline), um die Zusätzlichkeit der Auswirkungen im Vergleich zum Status quo zu gewährleisten.
3. Ausgewählter Ansatz (statisch / dynamisch)
 - Statische Betrachtung: Betrachtung nur der aktuellen Entscheidungssituation für eine Maßnahme.
 - Dynamische Betrachtung: Berücksichtigung des Planungsstandes einer Maßnahme (z.B. durch Aktualisierung der finanziellen Mittel o.ä.).
4. Methodisches Verfahren zur Ermittlung der (Brutto-)Auswirkungen der Maßnahme
 - Bestimmung des Wirkungsmodells.
 - Festlegung der Berechnungsmethodik zur Ermittlung der Bruttoauswirkungen (jeweils abhängig von den Besonderheiten einer Maßnahme).
5. Methodisches Verfahren zur Anpassung der Wirkung (Nettoeffekt der Maßnahme)
 - Effektbereinigung auf der Ebene einer Einzelmaßnahme (Netto 1): Bereinigung des Bruttoeffekts um Mitnahme-, Vorzieh-, Spillover-, Verzögerungs-, Struktur- und Rebound-Effekte.
 - Effektbereinigung auf der Ebene eines Maßnahmenbündels (Netto 2): Bereinigung des Netto-1-Effekts um Wechselwirkungen zwischen den einzelnen Maßnahmen.
6. Umgang mit Ungewissheiten
 - Verringerung der methodischen Unsicherheiten durch Anwendung einer einheitlichen und umfassenden Evaluierungsmethodik.
 - Transparente Dokumentation der verbleibenden Unsicherheiten (z. B. aufgrund von regulatorischen Unsicherheiten, Preisunsicherheiten, mangelnder Akzeptanz usw.).



Der rote Faden durch die Energiewende: Das Kopernikus-Projekt Ariadne führt durch einen gemeinsamen Lernprozess mit Politik, Wirtschaft und Gesellschaft, um Optionen zur Gestaltung der Energiewende zu erforschen und politischen Entscheidern wichtiges Orientierungswissen auf dem Weg zu einem klimaneutralen Deutschland bereitzustellen.

Folgen Sie dem Ariadnefaden:

 @AriadneProjekt

 Kopernikus-Projekt Ariadne

 ariadneprojekt.de

Mehr zu den Kopernikus-Projekten des BMBF auf kopernikus-projekte.de

Wer ist Ariadne? In der griechischen Mythologie gelang Theseus durch den Faden der Ariadne die sichere Navigation durch das Labyrinth des Minotaurus. Dies ist die Leitidee für das Energiewende-Projekt Ariadne im Konsortium von 27 wissenschaftlichen Partnern. Wir sind Ariadne:

adelphi | Brandenburgische Technische Universität Cottbus-Senftenberg (BTU) | Deutsches Institut für Wirtschaftsforschung (DIW) | Deutsches Zentrum für Luft- und Raumfahrt (DLR) | Ecologic Institute | Forschungsinstitut für Nachhaltigkeit – Helmholtz-Zentrum Potsdam (RIFS) | Fraunhofer Cluster of Excellence Integrated Energy Systems (CINES) | Friedrich-Alexander-Universität Erlangen-Nürnberg | Hertie School | ifo Institut – Leibniz-Institut für Wirtschaftsforschung an der Universität München | Institut der deutschen Wirtschaft Köln | Julius-Maximilian-Universität Würzburg | Mercator Research Institutes on Global Commons and Climate Change (MCC) | Öko-Institut | Potsdam-Institut für Klimafolgenforschung (PIK) | RWI – Leibniz-Institut für Wirtschaftsforschung | Stiftung Umweltenergierecht | Stiftung Wissenschaft und Politik | Technische Universität Berlin | Technische Universität Darmstadt | Technische Universität München | Universität Duisburg-Essen | Universität Greifswald | Universität Hamburg | Universität Potsdam | Universität Stuttgart – Institut für Energiewirtschaft und Rationelle Energieanwendung (IER) | ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung