

Stellungnahme zum Antrag der Fraktion der FDP

## Wirksame, digitale und transparente Familienleistungen – Die Evaluation von ehe- und familienpolitischen Leistungen als dauerhafter Prozess

Prof. Dr. Holger Bonin  
Forschungsinstitut zur Zukunft der Arbeit (IZA), Bonn, und Universität Kassel

7. Januar 2019

### 1. Ausgangslage

Nach der groß angelegten Evaluation der aktiven Arbeitsmarktpolitik im Umfeld der Hartz-Reformen war die Familienpolitik das zweite Maßnahmenfeld im Zuständigkeitsbereich des Bundes, auf dem die vorhandene Vielzahl von Leistungen einer umfassenden Bewertung unterzogen wurde. Einen wichtigen Anstoß dazu gab der Siebte Familienbericht, der einen Perspektivwechsel von Familienpolitik als sozialer Ausgleichspolitik hin zu Familienpolitik als nachhaltiger Unterstützung von Eltern und Kindern über den Lebenslauf unterstützte. Die Sachverständigenkommission legte in ihrem Bericht offen, dass eine auf Nachhaltigkeit ausgerichtete Familienpolitik nicht ohne Evaluation der eigenen Maßnahmen auskommen kann.

In der Folge beauftragte das BMFSJ gemeinsam mit dem BMF im Herbst 2009 die „Gesamtevaluation der ehe- und familienbezogenen Leistungen in Deutschland“. Dieses über vier Jahre laufende Großvorhaben hat durch eine vergleichende Bewertung von Effektivität und Effizienz von 13 zentralen Instrumenten der Familienförderung sowie ihrer Interaktionen umfassende Evidenz als Basis für eine sowohl zielorientierte als auch wirtschaftliche Politikgestaltung geliefert. Die Gesamtevaluation konzentrierte sich auf Geldleistungen, steuerrechtliche Maßnahmen, Sozialversicherungsleistungen und Realleistungen, die Eltern hauptsächlich während der Lebensphase erreichen, in der ihre Kinder heranwachsen. Die untersuchten Leistungen deckten rund 75 Prozent des finanziellen Gesamtvolumens der in einer vom BMFSJ für das Jahr 2010 erstellten Bestandsaufnahme von gut 150 familienbezogenen Leistungen und Maßnahmen des Staates ab. Der Abdeckungsgrad war sogar noch höher, wenn man aus diesem Tableau die Leistungen der Jugendhilfe (ohne Betreuung) nach SGB VIII, die Leistungen der Bildungsförderung für Studierende und Schüler nach BAföG sowie die ehe- und kinderbezogenen Ansprüche für Beamtinnen und Beamte nach BbesG und BeamtVG herausnimmt, die nur im weiteren Sinn als familienpolitische Leistungen zu werten sind. Die nicht abgedeckten Leistungen verteilen sich auf eine Vielzahl von häufig geringfügigen Einzelmaßnahmen und -leistungen für spezifische Zielgruppen, deren Inanspruchnahme sich in repräsentativen Haushaltsdaten kaum beobachten lässt.

Der Prozess der umfassend angelegten Bewertung der Familienförderung wirkt bis heute nach, da mit ihm die Voraussetzungen für eine stärkere Evidenzbasierung der Familienpolitik in verschiedener Hinsicht gestärkt wurden.

Erstens wurde im Vorfeld der Gesamtevaluation eine klare Festlegung auf Ziele erreicht, bei denen die Familienpolitik mit ihren Instrumenten Verbesserungen anstrebt: die wirtschaftliche Stabilität und soziale Teilhabe von Familien, die Vereinbarkeit von Familie und Beruf, die Förderung und das Wohlergehen von Kindern, die Realisierung von Kinderwünschen, sowie der Nachteilsausgleich zwischen den Familien. Diese für die Gesamtevaluation maßgeblichen Ziele haben normativen Charakter und lassen somit hinsichtlich Auswahl und Interpretation Raum für andere Sichtweisen und Kritik. Sie bleiben jedoch im aktuellen Diskurs über die Gestaltung der Familienpolitik in Deutschland wie in wissenschaftlichen Evaluationsstudien viel gebrauchte Benchmarks.

Zweitens hat die Gesamtevaluation die Datenbasis für eine evidenzbasierte Familienpolitik in Deutschland nachhaltig verbreitert. Der Aufwand zur Verbesserung der Dateninfrastruktur musste betrieben werden, weil die Inanspruchnahme von Leistungen im Bereich der Familienpolitik nur sehr unvollständig administrativ erfasst ist. Zudem sind familienpolitisch relevante Zielgrößen, vor allem die Förderung und das Wohlergehen von Kindern betreffend, in amtlich erhobenen Daten nicht ausreichend abgebildet. Es ist gelungen, die für die Zwecke der Gesamtevaluation angestoßenen Längsschnittbefragungen nach deren Ende in das Socio-oekonomische Panel (SOEP) zu integrieren und damit weiterzuführen.

Drittens wurden im Rahmen der Gesamtevaluation verschiedene verhaltensfundierte Mikrosimulationsmodelle weiterentwickelt oder neu aufgebaut. Solche Modelle werden für eine umfassende Bewertung der Wirksamkeit der Familienpolitik benötigt, um auch die Effekte der komplexen Interaktionen zwischen den einzelnen ehe- und familienbezogenen Leistungen und Maßnahmen innerhalb des Steuer- und Transfersystems abzubilden. Die mit der Gesamtevaluation geschaffenen Modellinfrastrukturen sind weiterhin nutzbar und wurden wiederholt zur Vorabbewertung von Veränderungen am bestehenden System der familienpolitischen Leistungen eingesetzt. So haben Bonin et al. (2018) im Vorfeld der geplanten Weiterentwicklung des Kinderzuschlags im Auftrag der von der Prognos AG geführten Geschäftsstelle des Kompetenzbüros Wirksame Familienpolitik die möglichen Arbeitsangebotseffekte unterschiedlicher Reformvarianten vorausberechnet.

Viertens haben die Erfahrungen mit der umfassend angelegten und komplexen Gesamtevaluation zur Entwicklung einer Evaluationskultur im Bereich der Familienpolitik des Bundes beigetragen. Ein Beleg dafür sind die inzwischen zum Standard gehörenden Evaluationsklauseln bei Einführung oder substanziellen Veränderung von Familienleistungen des Bundes. Die gewachsene Evaluationskultur zeigt sich aber auch an einer intensivierten Kommunikation zwischen Wissenschaft, Administration und Politik über Fragen einer zielgerechten Gestaltung der deutschen Familienpolitik. In den an der Gesamtevaluation beteiligten Ressorts ist ein solides Verständnis der im Kontext der Familienpolitik gebrauchten Methoden der Evaluierung, ihrer Voraussetzungen und Limitationen entstanden. Das gewachsene Verständnis unterstützt die Formulierung wissenschaftlich anspruchsvoller spezifischer Untersuchungsvorhaben, die bei Bedarf an handlungsorientierender wissenschaftlicher Expertise und empirischer Evidenz gezielt abgerufen werden.

## **2. Bewertung des Antrags der FDP-Fraktion**

### **2.1. Verstetigung der Gesamtevaluation**

Eine zentrale Forderung im Antrag der FDP-Fraktion ist, die Evaluation aller ehe- und familienpolitischen Leistungen als dauerhaften Prozess zu etablieren. Eine im regelmäßigen Turnus durchgeführte Gesamtevaluation der Familienleistungen erscheint allerdings nur insoweit sinnvoll, wie der zu erwartende Nutzen in einer vertretbaren Relation zu den damit verbundenen Kosten steht. Dieses Kriterium führt direkt zu der Frage, ob sich bei einer Wiederholung des Untersuchungsprogramms der Gesamtevaluation neue Evidenz gewinnen ließe, die für die Weiterentwicklung der Familienpolitik hilfreich wäre. Dagegen spricht zunächst, dass es zwischenzeitlich keine fundamentalen Fortschritte bei der für die Evaluation der Familienpolitik nutzbaren Methoden und Datengrundlagen gegeben hat. Gegen substantielle Erkenntnisgewinne spricht zudem, dass trotz gewisser Neuerungen in der Familienpolitik Funktionsweise und Struktur des Systems der ehe- und familienbezogenen Leistungen seit dem Abschluss der Gesamtevaluation dem Grunde nach unverändert geblieben sind. Somit erscheint die Wahrscheinlichkeit hoch, dass ein baldiges Update der Gesamtevaluation die bereits vorhandene Evidenz zumindest qualitativ größtenteils lediglich doppeln würde.

Andererseits könnte ein Wandel der Rahmenbedingungen die Effektivität und Effizienz der Familienleistungen beeinflussen, auch wenn sich diese selbst gar nicht verändern. Erstens sind die Bedingungen am Arbeitsmarkt heute deutlich andere als im durch die Gesamtevaluation abgedeckten Untersuchungszeitraum. So könnten der anhaltende Beschäftigungsaufbau und steigende Reallöhne die Voraussetzungen für eine gute Vereinbarkeit von Familie und Beruf unabhängig von der Gestaltung der Familienleistungen verbessert haben. Zweitens haben sich die demografischen Rahmenbedingungen verändert, nicht nur durch den fortschreitenden natürlichen Alterungsprozess in der einheimischen Bevölkerung, sondern auch durch hohe Wanderungsgewinne im Kontext von Arbeits- und Fluchtmigration. Damit erreicht die Familienpolitik heute neue Zielgruppen und entfaltet bei diesen möglicherweise eine andere Wirksamkeit als in der früher beobachteten Bevölkerung. Drittens kommen in Deutschland zunehmend Kohorten in die Familienphase, die andere Vorstellungen etwa im Hinblick auf den Wert von Erwerbsarbeit oder die gleichberechtigte Teilung von Erwerbs- und Sorgearbeit mitbringen. Veränderte Präferenzen können bedeuten, dass die den Berechnungen der Gesamtevaluation zugrundeliegenden Verhaltensparameter nicht mehr gültig sind. Entsprechend können sich auch die abzuschätzenden Wirkungen gegebener Familienleistungen verändern.

Diese Treiber von möglichen Veränderungen der Wirkungen bestehender Familienleistungen dürften allerdings kaum stark genug sein, um damit schon nach wenigen Jahren eine vollständige Neuauflage des aufwändigen Untersuchungsprogramms der Gesamtevaluation zu rechtfertigen. Adäquat erscheint vielmehr die im Vergleich viel weniger aufwändige Verstetigung einer Kern-Evaluation der Familienleistungen im Sinne eines Monitoring. Damit ist gemeint, die Wirkungen der zentralen ehe- und familienpolitischen Leistungen in regelmäßigem Turnus – etwa in einer Frequenz von vier Jahren – mittels eines aus der Gesamtevaluation zu übernehmenden Basismodells zu untersuchen, das mit den jeweils neuesten verfügbaren Daten gespeist und methodisch auf den jeweils aktuellsten Stand des Wissenschaft gebracht wird.

Mit dieser Vorgehensweise ließen sich konsistente Zeitreihen von Evaluationsergebnissen gewinnen. Die bei einem verstetigten Monitoring in den Beobachtungsreihen erkennbar werdenden Trends und Strukturbrüche könnten Hinweise auf relevante Veränderungen in den wirtschaftlichen, demografischen und gesellschaftlichen Rahmenbedingungen liefern, die eine Adjustierung einer wirkungs- und effizienzorientiert ausgerichteter Familienleistungen erforderlich machen.

## **2.2. Priorisierung der familienpolitischen Ziele**

Evaluationsforschung im öffentlichen Auftrag benötigt stets klare Vorgaben für die Ziele, anhand derer die Effektivität der zu untersuchenden Leistungen bewertet werden soll. Empirische Wirkungsforschung als positive wissenschaftliche Analyse kann sich diese Zielvorgaben nicht selbst geben. Ihr obliegt lediglich die Aufgabe, die von der Politik verfolgten normativen Ziele inhaltlich angemessen und transparent in messbare Zielvariablen zu überführen. Der von der Politik vorgegebene Katalog von Zielen muss in sich widerspruchsfrei sein. Damit ist gemeint, dass sich die zu untersuchenden Zieldimensionen nicht grundsätzlich gegenläufig zueinander verhalten dürfen. Ein Beispiel hierfür wären die im Kontext der Familienpolitik diskutierten Ziele der Vereinbarkeit von Familie und Beruf, zu beurteilen etwa an den Erwerbsquoten von Eltern, und der Wahlfreiheit, zu messen etwa an den Raten der Nicht-Beteiligung von Eltern am Erwerbsleben.

Die in der Gesamtevaluation verfolgten Ziele sind in sich widerspruchsfrei. Ansonsten könnte es keine Leistungen geben, die im Hinblick auf alle vorgegebenen Ziele positiv wirken, wie es gemäß den Evaluationsergebnissen etwa bei der nicht kostendeckenden öffentlichen Bereitstellung der Kindertagesbetreuung der Fall ist. Andererseits garantieren in sich konsistente Zielvorgaben nicht, dass eine Leistung bezogen auf einzelne Ziele in die gewünschte Richtung und zugleich bezogen auf andere Ziele in die ungewünschte Richtung wirkt. Einen Konflikt in diesem Sinn gibt es gemäß den Ergebnissen der Gesamtevaluation zum Beispiel beim Ehegattensplitting, das zumindest kurzfristig zur wirtschaftlichen Stabilität von Familien beiträgt, zugleich aber dem Vereinbarkeitsziel entgegenläuft. Bei einer Priorisierung des Zielkatalogs derart, dass eine Hierarchisierung der darin enthaltenen Einzelziele vorgegeben wird, würde man also beim des Ehegattensplitting je nach angelegtem Gewichtungsschema zu unterschiedlichen Bewertungen kommen. Eine umfassende vergleichende Evaluation dient allerdings eher einer Priorisierung der untersuchten Leistungen in dem Sinn, wirksamen Instrumenten, die nicht zu Zielkonflikten führen, im Rahmen der Politikgestaltung ein höheres Gewicht zu geben als konfliktbehafteten Instrumenten.

Aus diesen Feststellungen folgt: Sofern die im Antrag der FDP-Fraktion erhobene Forderung, bei einer Evaluation von ehe- und familienbezogenen Leistungen eine Priorisierung der familienpolitischen Ziele vorzunehmen, bedeutet, diese Ziele überhaupt konkret und in sich widerspruchsfrei zu benennen, wird damit lediglich auf eine basale Voraussetzung von Vorhaben der Programmevaluation zur Politikberatung hingewiesen. Die Leistungsbeschreibungen zu Evaluationsvorhaben im Zuständigkeitsbereich des BMFSFJ enthalten schon jetzt regelmäßig entsprechende Vorgaben. Bei einer Wiederholung der Gesamtevaluation – oder dem vorzuziehenden weniger umfassenden Monitoring in regelmäßigem Turnus – wäre der Katalog der Zielgrößen zu adjustieren, um gegebenenfalls zwischenzeitlich veränderten Zielvorstellungen der Familienpolitik Rechnung zu tragen. Allerdings erscheinen die für im Vorfeld der Gesamtevaluation festgelegten Zieldimensionen als Orientierungspunkte für die Gestaltung der Familienpolitik des Bundes immer

noch aktuell. Sollte die Forderung nach einer Priorisierung der familienpolitischen Ziele dagegen auf eine Hierarchisierung beziehungsweise Gewichtung verschiedener Ergebnisdimensionen abzielen, wäre diese kritischer zu sehen. Diese Art der Priorisierung würde nämlich Spielräume öffnen, im Hinblick auf einzelne Dimensionen negative Evaluationsergebnisse bei der politischen Bewertung durch Nachjustierung der Gewichte zu akkommodieren.

### **2.3. Inhalte einer weiteren umfassenden Evaluation**

Der Antrag der FDP-Fraktion enthält verschiedene Forderungen betreffend Inhalt und Fragestellungen einer neuen umfassenden Evaluation der Familienleistungen, die über das Programm der zurückliegenden Gesamtevaluation deutlich hinausgehen. Diese Erweiterungsforderungen verweisen zwar durchaus auf inhaltlich relevante Themenstellungen. Wie die folgenden näheren Ausführungen zeigen, erscheint deren adäquate Bearbeitung allein schon wegen der dafür unzureichenden Datengrundlagen auf absehbare Zeit nicht adäquat machbar.

Erstens wird angeregt, in die Wirkungsanalysen weitere Familienleistungen des Bundes einzubeziehen, die in der vorliegenden Gesamtevaluation der ehe- und familienbezogenen Leistungen nicht behandelt wurden. Dabei geht es allerdings um eine Vielzahl von Einzelleistungen, die häufig nur sehr eng begrenzte Zielgruppen erreichen, und zu deren Inanspruchnahme derzeit kaum Daten verfügbar sind, wie sie für wissenschaftlich angemessene empirische Wirkungsanalysen benötigt werden. Darum erscheint eine Evaluation dieser Bundesleistungen kurzfristig nicht umsetzbar, worauf schon die der Gesamtevaluation vorgeschaltete Machbarkeitsstudie hingewiesen hat.

Prinzipiell wäre es zwar denkbar, auf längere Sicht die notwendigen Datengrundlagen für eine mehr in die Breite der Familienleistungen des Bundes gehende Evaluation zu legen. Zu diesem Zweck müsste man allerdings wegen der spezifischen Zielgruppen Befragungen mit großen Teilnehmerzahlen durchführen, heterogene administrative Datenquellen zusammenführen und harmonisieren, oder größere administrative Datenlücken schließen. Jeder dieser Ansätze ist mit ganz erheblichem Aufwand verbunden, und angesichts des recht geringen Gewichts der nicht in der Gesamtevaluation abgebildeten Familienleistungen des Bundes erscheint es ziemlich fraglich, dass dieser Aufwand in einem angemessenen Verhältnis zum damit für die Politik erzielbaren Erkenntnisgewinn stehen würde.

Zweitens fordert die FDP-Fraktion, in eine umfassendere Evaluation die Interaktionen zwischen Bundesleistungen und Familienleistungen der Länder- und kommunalen Ebene ausdrücklich einzubeziehen. In den Ergebnissen der Gesamtevaluation waren die Effekte dieser Leistungen, die Wirksamkeit und Effizienz der untersuchten Bundesleistungen verstärken oder hemmen können, lediglich unterschwellig enthalten. Bei den mit Kontrollgruppenverfahren identifizierten Wirkungen von Familienleistungen des Bundes handelt es sich um Nettoeffekte; diese nehmen indirekt auch solche Veränderungen der Zielgrößen auf, die von nicht beobachteten Leistungen auf Länder- oder kommunaler Ebene ausgehen, soweit deren Inanspruchnahme mit der Inanspruchnahme der untersuchten Bundesleistungen systematisch korreliert ist. Bei den mit Mikrosimulationsmodellen durchgeführten Analysen, die in der Gesamtevaluation gegenüber den Kontrollgruppenanalysen stark überwiegen, können sich in den zur Parametrisierung der Modelle geschätzten Verhaltensparametern implizit auch auf Länder oder kommunaler Ebene bestimmte Kontextbedingungen spiegeln. Für die von der FDP-Fraktion geforderte tiefer gehende Untersuchung müssten die

Schnittstellen zwischen Familienleistungen der Bundes-, Länder- und kommunalen Ebene allerdings ausdrücklich modelliert werden.

Momentan stehen entsprechend elaborierte Steuer-Transfer-Modelle für Deutschland nicht zur Verfügung. Ihre Entwicklung scheitert zum einen an der Komplexität des Geflechts föderaler Strukturen. Diese bringt sowohl auf der Länderebene als auch innerhalb einzelner Bundesländer auf kommunaler Ebene stark unterschiedliche Ausprägungen der monetären und nicht-monetären Unterstützungsangebote für Familien hervor, die Bundesleistungen ergänzen oder substituieren können. Zum anderen fehlt ein über länder- oder kommunale Grenzen hinweg einheitliches Informationserfassungssystem. Deswegen sind umfassende Daten zur Inanspruchnahme der diversen unterhalb der Bundesebene organisierten Familienleistungen, wie sie für umfassendere Wirkungsanalysen gebraucht würden, nicht verfügbar. Es erscheint auch praktisch ausgeschlossen, eine entsprechende Datenbasis eigens für den Zweck einer Evaluation der Familienleistungen aufzubauen. Ebenfalls ist auf absehbare Zeit nicht damit zu rechnen, dass eine hinreichende Harmonisierung der dezentral betriebenen Datenerfassung erreicht werden kann, sodass eine vertikale und horizontale föderale Strukturen abbildende umfassende Evaluation von Familienleistungen auf administrativen Prozessinformationen aufbauen könnte.

Machbar erscheint angesichts dieser Hürden allenfalls das viel moderatere Vorhaben einer qualitativ ausgerichteten systematischen Aufarbeitung der rechtlichen und praktischen Schnittstellenprobleme, die sich beim Zusammenwirken von Familienleistungen der Bundes-, Landes- und kommunalen Ebene ergeben können. Eine solche Untersuchung würde an die im Rahmen der Gesamtevaluation ausgeführte Untersuchung zu Schnittstellen im Sozial-, Steuer- und Unterhaltsrecht anknüpfen, die herausgearbeitet hat, dass sich für Familien durch fehlende Abstimmung vieler Einzelregelungen unbeabsichtigte Wirkungen ergeben können, die den Gesetzeszwecken und familienpolitischen Zielen zuwiderlaufen. Dabei dürfte es, um den Zweck der besseren Information über die Rahmenbedingungen einer wirksamen und effizienten Familienpolitik des Bundes zu erfüllen, hinreichend sein, im Rahmen einer solchen Studie typische Schnittstellenprobleme anhand exemplarischer Konflikte zwischen Leistungen der Bundes-, Landes- und kommunalen Ebene herauszuarbeiten. Der beträchtliche Aufwand einer vollständigen Kartierung der Landschaft der Familienleistungen von Ländern und Kommunen dürfte dagegen in keinem angemessenen Verhältnis zum Informationsbedarf der Bundesebene stehen.

Drittens wird im Antrag der FDP-Fraktion gefordert, im Rahmen einer umfassenderen Evaluation über die Kriterien der Effektivität und Effizienz der Familienleistungen hinaus auch die Effizienz der administrativen Prozesse bei ihrer Verwaltung sowie bestehende Digitalisierungspotenziale zu adressieren. Mit diesen Erweiterungen würde die für die Gesamtevaluation maßgebliche analytische Ausrichtung auf die Quantifizierung von Wirkungen bei den Adressatinnen und Adressaten der Familienleistungen des Bundes aufgegeben. Die neue Fragestellung könnte im Gegensatz dazu nicht ohne qualitativ ausgerichtete Implementationsanalysen auskommen. Eine aussagekräftige Untersuchung von Digitalisierungspotenzialen würde zudem tiefgehende Analysen der dateninfrastrukturtechnischen, IT-technischen und datenschutzrechtlichen Voraussetzungen und Entwicklungsmöglichkeiten erfordern. Da auch Bürokratiekosten bzw. Einsparungspotenziale im Bereich der Verwaltung bewertet werden sollten, müssten zudem Daten erschlossen werden, die geeignet sind, diese zu quantifizieren.

Weil im Hinblick auf die Aufgabenstellungen, die mit einer Bewertung der Bereitstellungspraxis und Digitalisierungspotenziale bei Familienleistungen verbunden sind, nicht auf Vorerfahrungen aus der Gesamtevaluation zurückgegriffen werden kann, wäre zunächst ein dafür geeignetes methodisches Konzept zu entwickeln. Hierzu sollte der eigentlichen Untersuchung eine Machbarkeitsstudie vorgehen, um ein Untersuchungsprogramm zu konkretisieren, dass hohen wissenschaftlichen Anforderungen genügt, dem Informationsbedarf der Akteure gerecht wird und wirtschaftlich durchgeführt werden kann. Mit ähnlichen Zielsetzungen war auch der Gesamtevaluation eine Machbarkeitsstudie vorgelagert. Eine sorgfältige konzeptionelle Vorstudie erscheint vor Beauftragung einer Studie zu Effizienzspielräumen im Bereich der Verwaltung von Familienleistungen noch wichtiger. Dieser Gegenstand ist vergleichsweise schlecht erforscht, es muss ein komplexes Zuständigkeitsgeflecht durchdrungen werden, und die Datengrundlagen für eine empirisch adäquate Abschätzung von Bürokratiekosten sind erfahrungsgemäß reichlich schwach.

Bei diesen Ausgangsbedingungen erscheint eine vollständige Durchführung des komplexen Evaluationsauftrags, die Familienleistungen des Bundes nicht nur nach Wirksamkeit und Kosten-Nutzen-Verhältnis, sondern auch im Hinblick auf die Qualität der Verwaltungsstrukturen und die Digitalisierung von administrativen Leistungsprozessen zu untersuchen, innerhalb von zwei Jahren, wie sie von der FDP-Fraktion gefordert wird, nicht realistisch.