

Zusammenfassung des Gutachtens

für das Bundesministerium für Wirtschaft und Klimaschutz (BMWK)
zum Projekt Nr. 08/22

Potenzialschätzung und
Produktionslücken der
Bundesregierung –
Darstellung und Ver-
gleich von Vorschlägen
zur Anpassung der
Bestimmungsmethode



Kiel, den 02. Mai 2023

Institut für Weltwirtschaft
an der Universität Kiel

Im Rahmen der Konjunkturbereinigung in der nationalen Schuldenregel in Deutschland erfolgt die Bestimmung des gesamtwirtschaftlichen Produktionspotenzials nach dem Produktionsfunktionsansatz in Übereinstimmung mit dem nach dem Stabilitäts- und Wachstumspakt angewandten Verfahren (gemeinsame EU-Methode, EU-CAM). Dabei beschreibt das gesamtwirtschaftliche Produktionspotenzial die Wirtschaftsleistung in einer konjunkturellen Normallage entsprechend dem Grundgesetz. Hierfür ist eine Schätzung erforderlich, da diese nicht beobachtbar ist.

Im Koalitionsvertrag zur 20. Legislaturperiode wurde vereinbart, dass die Konjunkturbereinigung, unter anderem basierend auf den Erkenntnissen der vergangenen zehn Jahre, z. B. vor dem Hintergrund von systemischen Krisen, im Rahmen der nationalen Schuldenregel evaluiert und entsprechend der sich daraus ergebenden Bedarfe angepasst werden soll, ohne die grundgesetzliche Schuldenregel zu ändern.

Mit einer Auftaktkonferenz am 5. Juli 2022 wurde ein wissenschaftlich gestützter Beteiligungsprozess des Bundesministeriums für Wirtschaft und Klimaschutz (BMWK) gestartet, der ermöglichen soll, dass breite und tiefgehende Erkenntnisse zur Potenzialschätzung bei der Evaluierung berücksichtigt werden können. Die Auftaktkonferenz wurde von einem „Call for Proposals“ gefolgt, der zur Einreichung von Vorschlägen zur Potenzialschätzung einlud. Als Zwischenschritt fand am 24. August 2022 ein wissenschaftlicher Workshop statt, auf dem die Einreichenden die Möglichkeit hatten, ihre Vorschläge einem Fachpublikum vorzustellen. Anschließend bestand bis zum 15. September 2022 Zeit, die Vorschläge ggf. zu überarbeiten und schließlich eine operationalisierbare Form für die Berücksichtigung in dem vorliegenden wissenschaftlichen Gutachten einzureichen. Die Teilnahme am vorherigen Workshop war keine Vorbedingung für die Einreichung.

Der Kerngegenstand dieses Gutachtens besteht darin, die schließlich eingegangenen folgenden sieben Vorschläge empirisch vergleichend darzustellen und anhand von Echtzeitdaten zu evaluieren:

1. Christian Breuer: Arbeitslosigkeit als natürlicher Konjunkturindikator (zwei Ansätze: zum einen basierend auf einer fixen Zielarbeitslosenquote und zum anderen auf der NAWRU des Benchmark-Verfahrens)
2. Kai Carstensen, Felix Kießner, Thies Rossian: Faktormodell zur Bestimmung der TFP
3. Dezernat Zukunft: Vorschläge zur Projektion des Arbeitspotenzials und der Produktionsfunktion
4. Marcell Göttert und Timo Wollmershäuser: Umfragegestützte Schätzung der Produktionslücke
5. Thorsten Hild: Produktionspotenzial als Produkt aus Erwerbspersonen und Arbeitsproduktivität
6. Carsten-Patrick Meier: Modellierung der NAWRU im D*-Modell von Kiel Economics
7. Christian Proaño: F&E-Indikatoren zur Verbesserung der TFP-Schätzung

wobei der Vorschlag von Proaño und zwei Teilvarianten des Vorschlags von Breuer aus technischen Gründen nicht in der Echtzeitdatenanalyse umgesetzt werden können (Breuer: fehlende Datenverfügbarkeit, Proaño: fehlendes Fortschreibungsmodell für Projektionszeitraum).

Die Echtzeitdatenergebnisse der Vorschläge werden denen des aktuell praktizierten Verfahrens der Bundesregierung (Benchmark), das ebenfalls für den Echtzeitdatensatz angewendet wird, sowie diverser Modifikationen des Benchmark-Verfahrens gegenübergestellt. Neben den technischen Modifikationen (NAWRU-Anker, Geflüchtetenanpassung und Auswahl der Fortschreibungsmodelle) werden Varianten mit zusätzlicher Glättung betrachtet, angelehnt an einen Vorschlag der Deutschen Bundesbank zur Glättung der Steuereinnahmen. Schließlich werden zusätzlich zwei gängige univariate Zeitreihenverfahren als weitere Vergleichspunkte herangezogen. Sofern technisch möglich, werden die eingereichten Vorschläge und die Modifikationen anhand der Eigenschaften, u. a. hinsichtlich Revisionen und Prozyklizität, verglichen, die diese bei Anwendung auf Echtzeitdaten aufweisen. Der besondere Fokus liegt dabei auf den Ergebnissen für das jeweils folgende Jahr („t+1“), denn der Haushalt des Folgejahres wird im jeweils laufenden Jahr („t“) erstellt und an die unterjährigen Revisionen jeweils angepasst Als

Echtzeitdaten werden die Datenstände inklusive Kurzfristprognosen herangezogen, die die Bundesregierung in den vergangenen zehn Jahren seit Herbst 2012 in ihren Potenzialschätzungen genutzt hat. Bei einigen Vorschlägen kann aufgrund von Datenlimitationen die Echtzeitdatenanalyse nur eingeschränkt umgesetzt werden (vgl. Tabelle 1, S. 19). Zudem wird die Schätzung der Produktionslücke des jüngsten Datenstands (Herbstprojektion 2022) dargestellt und mit Umfragezahlen zur gesamtwirtschaftlichen Kapazitätsauslastung verglichen, die vielfach als alternatives Maß für die konjunkturelle Entwicklung und damit für Produktionslücken gesehen werden. Dadurch soll überprüft werden, ob die Verfahren grundsätzlich in der Lage sind, plausible Produktionslückenschätzungen im Kontext der Konjunkturbereinigung zu erzeugen. Eine Diskussion bzw. Bewertung der theoretischen Grundlagen der eingereichten Arbeiten erfolgt im Gutachten nicht, wobei darauf hinzuweisen ist, dass angesichts des relativ kleinen Echtzeitdatensatzes theoretische Überlegungen durchaus in die finale Bewertung der Vorschläge aufgenommen werden sollten. Auch eine juristische Prüfung der Vorschläge bezüglich der grundgesetzlichen Anforderungen ist nicht Bestandteil dieses Gutachtens.

Insgesamt können und sollten zur Beurteilung eines Potenzialschätzverfahrens im Kontext der Konjunkturbereinigung verschiedene Kriterien herangezogen werden. Aus Sicht der Planungssicherheit bzw. Stetigkeit der öffentlichen Haushalte ist es grundsätzlich vorteilhaft, wenn die Potenzialschätzung wenig revidiert wird und wenig prozyklisch ist.

Bei den Revisionseigenschaften zeigt sich in der quantitativen Darstellung, dass die *mittleren Revisionen* in vielen der Vorschläge und Methodenansätze relativ gering sind (vgl. Tabelle 6, S. 28). Bei der ergänzend bestimmten *Standardabweichung der Revisionen*, die misst, mit welchem Revisionsumfang unabhängig von der Richtung der Revisionen im Mittel zu rechnen ist, zeigen sich deutlichere Unterschiede. Während die Glättungs-Ansätze des Benchmark-Verfahrens und der Vorschlag von Dezernat Zukunft relativ gut abschneiden, weisen die Ansätze, die ausschließlich auf die Arbeitslosenquote fokussiert sind (Breuer und Hild), sehr starke Revisionen auf (Tabelle 7, S. 28). Wenig *prozyklisch* zeigen sich die Ansätze von Dezernat Zukunft und auch von Meier. Hier dürften insbesondere die Setzungen, die in diesen beiden Verfahren vorgenommen werden und die nicht von der aktuellen Prognose des Bruttoinlandsprodukts abhängen, das gute Abschneiden begünstigen. Auch die zusätzliche Glättung des Benchmark-Verfahrens reduziert deutlich die Prozyklizität (vgl. Tabelle 8, S. 29). Die Ansätze von Breuer und von Hild sind hingegen stark prozyklisch, was damit zusammenhängen dürfte, dass Produktivitätstrends nicht berücksichtigt werden. Starke Revisionen und hohe Prozyklizität weisen zudem der HP-Filter, aber auch der Ansatz von Göttert und Wollmershäuser sowie etwas schwächer der modifizierte Hamilton-Filter auf. Dies ist zunächst überraschend, da sowohl der Ansatz von Göttert und Wollmershäuser als auch der Hamilton-Filter für Vergangenheitsdaten nahezu revisionsfrei sind. Hier ist anzumerken, dass die Revisionsanalyse an dieser Stelle mit Blick auf die Haushaltsplanung auf zukünftige Werte des Bruttoinlandsprodukts fokussiert ist und somit Revisionen des Bruttoinlandsprodukts bzw. dessen Prognosen die Potenzialschätzung in diesen Ansätzen stark beeinflussen können.

Zu betonen ist generell, dass die Revisionsintensität des Potenzials alleine *kein hinreichendes Gütekriterium* ist. So wäre beispielsweise eine naive Regel wie die, das Potenzial stets auf das Bruttoinlandsprodukt von vor zehn Jahren zu setzen, komplett revisionsfrei, würde aber keine plausible Potenzial- bzw. plausible Produktionslückenschätzungen erbringen. Zusätzlich wurde daher als zentrales Kriterium geprüft, ob die resultierenden Produktionslücken ein *plausibles Konjunkturmuster* im Sinne von rein konjunkturell bedingten Abweichungen von der „Normallage“ ergeben. Hierzu wurde der Vergleich mit Umfragen zur Kapazitätsauslastung („CUBS-Indikator“) als Plausibilitätskriterium operationalisiert. Zudem sollte ein plausibles Verfahren in anerkannten großen Konjunkturkrisen, wie der Finanzkrise 2008/2009 oder der Corona-Krise 2020, größere Produktionslücken zeigen als in anderen, konjunkturell ruhigeren Zeiten.

Viele der Vorschläge liefern im Vergleich zu den Umfragen der Kapazitätsauslastung und mit Blick auf die großen Konjunkturkrisen insgesamt plausible Ergebnisse (vgl. Abbildungen 8 bis 14, S. 22 f.). Entsprechend ergibt sich eine hohe Korrelation der jeweils geschätzten Produktionslücken mit Umfragedaten zur Kapazitätsauslastung, die bei Göttert und Wollmershäuser besonders stark ausgeprägt ist (Tabelle 2, S. 24). Sehr auffällig sind jedoch die Ansätze von Breuer (mit einer fixen strukturellen Arbeitslosigkeit) und von Hild, die bei den geschätzten Produktionslücken kaum auf die großen Konjunkturkrisen reagieren, eine Korrelation mit den Daten zur Kapazitätsauslastung ist nicht gegeben. Ebenfalls sehr auffällig ist die Produktionslücke von Dezernat Zukunft, die etwa im konjunkturell eher unauffälligen Jahr 2015 betragsmäßig ähnlich hoch lag wie im Jahr der Corona-Krise 2020. Das Konjunkturbild ausweislich von Umfragen zur Kapazitätsauslastung für diese beiden Jahre steht im Widerspruch dazu; so war der CUBS-Indikator im Jahr 2015 leicht positiv, zeigte im Jahr 2020 hingegen den dritt-niedrigsten Wert seiner Datenhistorie. Die Korrelation ist bei Dezernat Zukunft entsprechend nur sehr schwach ausgeprägt. Augenfällig ist zudem, dass die Produktionslücken für Breuer, Hild und Dezernat Zukunft trendbehaftet wirken und eine dauerhafte Unterauslastung, teilweise über Jahrzehnte hinweg, anzeigen. Die konjunkturell zulässige Neuverschuldung hätte sich bei Anwendung dieser Ansätze in der Tendenz zurückgebildet, wenn auch teilweise auf hohem Niveau. Für diese Ansätze ergeben sich deutlich negative mittlere Produktionslücken, auch in Echtzeit (vgl. Tabelle 4, S. 25). Nach den vorliegenden Ergebnissen kann nicht davon ausgegangen werden, dass bei den Ansätzen von Breuer, Hild und Dezernat Zukunft die Produktionslücke im Mittel auch über einen längeren Schätzzeitraum (nahe) Null ist, auch nicht für Echtzeitwerte. Damit verletzen diese Ansätze ein mögliches Symmetriekriterium. Der Ansatz von Meier weist zwar ein deutlich zyklisches Muster auf, der für die Konjunktur in der Finanzkrise jedoch die gleiche Produktionslücke ausgibt wie in den frühen 2000er Jahren über einen längeren Zeitraum; zudem ist die negative Lücke am aktuellen Rand auffällig groß.

Das Benchmark-Verfahren selbst ist in gewissem Umfang revisionsanfällig und prozyklisch, aber weniger stark als einige der untersuchten eingereichten Vorschläge und liefert bezogen auf die Umfragen zur Kapazitätsauslastung ein plausibles Konjunkturbild. Durch zusätzliche Glättung der Ergebnisse des Benchmark-Verfahrens kann die Revisionsintensität und Prozyklizität deutlich gemindert werden. Durch die Anbindung an das Benchmark-Verfahren bleibt dabei die wünschenswerte Eigenschaft des Benchmark-Verfahrens bezüglich einer nahe an Null liegenden durchschnittlichen Produktionslücke weitgehend erhalten. Die Verfahren zur Glättung des Produktionspotenzials erscheinen gerade zu Beginn von Krisen durchaus attraktiv, weil sie dann tendenziell größere fiskalische Spielräume erlaubten als das Benchmark-Verfahren selbst. Allerdings dürften sie in Aufschwüngen als restriktiv empfunden werden. Es kann zudem nicht ausgeschlossen werden, dass die zusätzliche Glättung zwar die Planbarkeit erhöht, aber zumindest zeitweise zu unplausiblen Produktionslückenschätzungen führt.

Die *technische Umsetzbarkeit* ist bei allen Verfahren nicht erheblich aufwändiger als beim Benchmark-Verfahren. Teilweise greifen Ansätze in weiten Teilen auf dieses zurück und erweitern es punktuell, was mit zusätzlichen Datenanforderungen und einer Komplexitätszunahme einhergeht (z.B. Carstensen et al.). Technisch sehr einfach sind die univariaten Zeitreihenverfahren, der Ansatz von Göttert und Wollmershäuser sowie die Ansätze von Hild und von Breuer. Allerdings ist diesen Ansätzen inhärent, dass sie keine eigene Orientierung für die Mittelfristprognose bieten können. Sollten diese Verfahren Anwendung finden, bräuchte es zusätzlich Techniken dafür, Projektionen für die mittelfristige Entwicklung abzuleiten. In den Berechnungen der Echtzeitdaten-Analyse in diesem Gutachten wurde für einige dieser Ansätze daher hilfsweise auf (Teile der) Mittelfrist-Ergebnisse des Benchmark-Verfahrens zurückgegriffen (Breuer, Hild) oder auf eine Betrachtung des Mittelfrist-Zeitraums verzichtet (Göttert und Wollmershäuser). Auch der Ansatz von Dezernat Zukunft ist auf Teile der Ergebnisse des Benchmark-Verfahrens angewiesen. Eine technische Herausforderung bietet unter Umständen das Verfahren von Meier. Die Parameterschätzungen fußen auf der einen Seite auf einem sehr langen Stützzeitraum, auf der anderen Seite werden explizit Strukturbrüche berücksichtigt. Dies bedeutet, dass regelmäßig Strukturbruchtests in Echtzeit angewendet werden müssen. Typischerweise bedarf es eines gewissen Stützzeitraums vor

und auch nach dem Bruchpunkt, um diesen mit hoher Sicherheit zu identifizieren. Die Gefahr besteht, dass bei der neuen Identifikation eines zusätzlichen Strukturbruchs das Konjunkturbild der vorausgegangenen Jahre erheblichen Revisionen unterworfen werden muss. In der hier durchgeführten Analyse hat sich zwar kein zusätzlicher Strukturbruch in Echtzeit gefunden. Ein schwer zu quantifizierendes Revisionsrisiko durch zusätzliche Strukturbrüche besteht allerdings bei diesem Verfahren.

Im Rahmen des Beteiligungsprozesses zur Potenzialschätzung des BMWK haben sich verschiedene Teilnehmerinnen und Teilnehmer zu relevanten Kriterien für die Beurteilung der Potenzialschätzverfahren geäußert. Wesentliche Aspekte sind demnach der Einfluss des Potenzialverfahrens auf die Planbarkeit der öffentlichen Haushalte, Transparenz und Nachvollziehbarkeit. Auch wurde mehrfach betont, dass eine Abstimmung mit dem Vorgehen auf europäischer Ebene wünschenswert sei. Einen Dissens unter den Teilnehmenden gab es bezüglich der Frage, ob die Produktionslücke im Mittel in etwa ausgeglichen sein sollte. Während viele Wortbeiträge betonten, dass dies grundgesetzlich intendiert und für die langfristige Steuerbarkeit des Verfahrens mit Blick auf die Tragfähigkeit der öffentlichen Finanzen notwendig sei, gab es zugleich Stimmen, die die Frage makroökonomischer Rückwirkungen als bedeutender einstufen.

Letztlich ist es Aufgabe der Regierung und des Gesetzgebers, im Rahmen der grundgesetzlichen Möglichkeiten zu entscheiden, welche der Kriterien und Aspekte besonders stark zu gewichten sind, um sich für oder gegen ein Konjunkturbereinigungsverfahren zu entscheiden. Dieses Gutachten liefert hierzu eine hilfreiche Entscheidungsgrundlage, für sich genommen jedoch keine abschließende Antwort auf die Frage, welches Verfahren vorzuziehen ist.