

RESEARCH REPORT SERIES

IZA Research Report No. 134

Machbarkeitsstudie für eine Untersuchung der sozialen Lage von Selbstständigen

Holger Bonin (IZA)

Annabelle Krause-Pilatus (IZA)

Ulf Rinne (IZA)

NOVEMBER 2022



Bundesministerium
für Arbeit und Soziales

FORSCHUNGSBERICHT 608

Machbarkeitsstudie für eine Untersuchung der sozialen Lage von Selbstständigen

November 2022

ISSN 0174-4992

Machbarkeitsstudie für eine Untersuchung der sozialen Lage von Selbstständigen

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

Forschungsinstitut zur Zukunft der Arbeit GmbH (IZA)
Schaumburg-Lippe-Straße 5–9
53113 Bonn

Holger Bonin (IZA)
Annabelle Krause-Pilatus (IZA)
Ulf Rinne (IZA)

August 2022

Erstellt im Auftrag des Bundesministeriums für Arbeit und Soziales.

Die Durchführung der Untersuchungen sowie die Schlussfolgerungen aus den Untersuchungen sind von den Auftragnehmern in eigener wissenschaftlicher Verantwortung vorgenommen worden. Das Bundesministerium für Arbeit und Soziales übernimmt insbesondere keine Gewähr für die Richtigkeit, Genauigkeit und Vollständigkeit der Untersuchungen.

Kurzbeschreibung

Das Bundesministerium für Arbeit und Soziales plant ein Forschungsprojekt mit dem Ziel, die sozialen Lagen von Selbstständigen, die sich daraus ergebenden besonderen Schutzbedarfe sowie die zu ihrer Erfüllung zielführenden sozialpolitischen Maßnahmen statistisch fundiert zu beschreiben. Dafür müssen Daten, welche die Grundgesamtheit der selbstständig tätigen Personen in Deutschland repräsentativ abbilden, erhoben werden. Die Expertise zeigt im Sinne einer Machbarkeitsstudie, wie ein zielführendes und wirtschaftliches Design für dieses Vorhaben aussehen könnte. Sie arbeitet die Vorteile der Anlage einer spezifischen Datenbasis gegenüber der Nutzung vorhandener Datenquellen heraus und erörtert die Möglichkeiten der praktischen Umsetzung einer Querschnittsbefragung, insbesondere Feldzugang und Stichprobendesign, sowie Fragenprogramm und Erhebungsmodus.

Abstract

The Federal Ministry of Labour and Social Affairs seeks to undertake a research project that aims at solid statistical description of the social situation of the self-employed in Germany, and the resulting protection requirements. For this purpose, it is necessary to gather data representing the universe of individuals who solely, partially or occasionally engage in self-employed work. This expertise shows, in the form of a feasibility study, how an effective and efficient design of the planned research project may look like given that a universal register of the target population does not exist. It demonstrates the advantages of constructing a specific data set, rather than using pre-existing data sets that are representative of the general population. It furthermore discusses the practical implementation of a specific cross-sectional survey, in particular the possibilities for access to the field and sampling design, the necessary content of the survey, and appropriate survey modes.

Inhaltsverzeichnis

Tabellenverzeichnis	7
Abbildungsverzeichnis	7
Abkürzungsverzeichnis	8
Zusammenfassung	10
1. Einleitung	12
2. Datenquellen und Studien mit Bezug zur sozialen Lage von Selbstständigen	18
2.1 Sozio-oekonomisches Panel	18
2.2 Panel Arbeitsmarkt und soziale Sicherung	19
2.3 Mikrozensus	20
2.4 HBS-Erwerbstätigenbefragung	21
2.5 IAB-Studien	22
2.6 Spezielle Befragungen von Selbstständigen	22
2.7 Daten zu Absicherung und Lebenslagen im Alter	24
3. Möglichkeiten für die Anlage einer neuen Erhebung bei Selbstständigen	26
3.1 Integration in laufende Befragungen	26
3.2 Grundgesamtheit und Zugangswege zur Zielgruppe	28
3.2.1 Unternehmensdatenbanken	28
3.2.2 Bevölkerungsrepräsentative Zufallsstichprobe	30
3.2.3 Selektive Zugangswege	33
3.3 Erhebungsverfahren	34
4. Befragungsinhalte	39
4.1 Basismerkmale auf Individual- und Haushaltsebene	40
4.2 Erwerbssituation und Charakter der selbstständigen Erwerbstätigkeit	41
4.3 Einkommen	43
4.4 Vermögen	44
4.5 Versicherung und Absicherung gegen zentrale Risiken	46
4.5.1 Alterseinkommen	46
4.5.2 Einkommensausfall	47
4.5.3 Gesundheit und Pflege	48
4.6 Erwerbsbiografie	48

4.7	Subjektive Einschätzungen, Motivationen und Haltungen sowie Persönlichkeitsmerkmale und Zufriedenheit	50
5.	Durchführung des Vorhabens	52
5.1	Einleitung	52
5.2	Erhebungsmodus	54
5.3	Abgrenzung der Zielgruppe	54
5.4	Umfang und Strukturierung der Nettostichprobe	56
5.5	Feldzugang	58
5.5.1	Einleitung	58
5.5.2	Aufbau einer dualen Stichprobe	59
5.5.3	Alternativen	61
5.6	Kostenabschätzungen	62
5.6.1	Aufbau einer dualen Stichprobe	63
5.6.2	Alternativen	65
5.7	Zeitbedarf	67
5.8	Entwicklungsperspektiven	68
	Literaturverzeichnis	71

Tabellenverzeichnis

Tabelle 1.1	Indikatoren zur Messung der Dimensionen von prekärer Beschäftigung auf der Personenebene	15
Tabelle 1.2	Indikatoren zur Messung der Dimensionen von prekären Lagen auf der Haushaltsebene	16
Tabelle 5.1	Geschätzter Kostenaufwand (netto) – Aufbau einer dualen Stichprobe	64
Tabelle 5.2	Geschätzter Kostenaufwand (netto) – Bevölkerungsrepräsentative Erhebung mit um 20 Prozent vergrößerter Nettostichprobe	65
Tabelle 5.3	Geschätzter Kostenaufwand (netto) – Bevölkerungsrepräsentative Erhebung mit Ergänzungsstichprobe aus Unternehmensdatenbank	66

Abbildungsverzeichnis

Abbildung 3.1	Schema eines sparsamen Screening-Designs	32
Abbildung 5.1	Aufbau einer dualen Stichprobe	60

Abkürzungsverzeichnis

ADM	Arbeitsgemeinschaft deutscher Marktforschungsinstitute
AKE-EU	Arbeitskräfteerhebung der Europäischen Union
ALV	Arbeitslosenversicherung
ASID	Studie „Alterssicherung in Deutschland“
BA	Bundesagentur für Arbeit
BAföG	Bundesausbildungsförderungsgesetz
BAMF	Bundesamt für Migration und Flüchtlinge
BAV	Betriebliche Altersvorsorge
BMAS	Bundesministerium für Arbeit und Soziales
BMWK	Bundesministerium für Wirtschaft und Klimaschutz
CAPI	computergestützte persönliche Interviews
CATI	computergestützte telefonische Interviews
CAWI	computerassistierte Web-Interviews
DSGVO	Datenschutz-Grundverordnung
EVS	Einkommens- und Verbrauchsstichprobe
GKV	Gesetzliche Krankenversicherung
GR	Gesetzliche Rentenversicherung
HBS	Hans-Böckler-Stiftung
IAB	Institut für Arbeitsmarkt- und Berufsforschung
IEB	Integrierte Erwerbsbiografien
ifo	Leibniz-Institut für Wirtschaftsforschung an der Universität München
IZA	Forschungsinstitut zur Zukunft der Arbeit
KfW	Kreditanstalt für Wiederaufbau
MIP	Mannheimer Innovationspanel
MUP	Mannheimer Unternehmenspanel
PAPI	Paper-and-Pencil Interviewing
PASS	Panel Arbeitsmarkt und soziale Sicherung
PKV	Private Krankenversicherung
RLD	Randomized Last Digit
SAVE	Studie „Sparen und AltersVorsorgE in Deutschland“
SGB II	Sozialgesetzbuch Zweites Buch
SGB III	Sozialgesetzbuch Drittes Buch
SGB X	Sozialgesetzbuch Zehntes Buch

SOEP	Sozio-oekonomisches Panel
SOEP-CoV	Studie „Sozio-ökonomische Faktoren und Folgen der Verbreitung des Coronavirus in Deutschland“ des Sozio-oekonomischen Panels
SOEP-IS	SOEP-Innovations-Stichprobe
VGR	Volkswirtschaftliche Gesamtrechnung
VGSD	Verband der Gründer und Selbstständigen Deutschland
WSI	Wirtschafts- und Sozialwissenschaftliches Institut in der Hans-Böckler-Stiftung
ZEW	Leibniz-Zentrum für Europäische Wirtschaftsforschung

Zusammenfassung

Die sozialen Lagen von Selbstständigen, die sich daraus ergebenden besonderen Schutzbedarfe sowie die zu ihrer Erfüllung zielführenden sozialpolitischen Maßnahmen lassen sich für Deutschland bislang nur unzureichend statistisch fundiert beschreiben. Das Bundesministerium für Arbeit und Soziales (BMAS) plant daher ein Forschungsprojekt mit dem Ziel, dafür geeignete Daten zu gewinnen. Die Studie soll die Grundgesamtheit der selbstständigen Personen im aktuellen Querschnitt repräsentativ abbilden und auch belastbare Aussagen zu Inzidenz und Folgen hybrider Arbeitsverhältnisse ermöglichen. Inhaltlich soll ein Fokus auf der Analyse prekärer Selbstständigkeit liegen. Prekarität wird dabei als multidimensionaler Sachverhalt verstanden, der sowohl verschiedene Kennzeichen prekärer Beschäftigung auf der individuellen Ebene als auch prekärer sozialer und wirtschaftlicher Lagen auf der Haushaltsebene umfasst. Die Expertise zeigt im Sinne einer Machbarkeitsstudie, wie ein zielführendes und wirtschaftliches Design für dieses Vorhaben aussehen könnte.

Ein Literaturüberblick zeigt, dass in Deutschland bereits bevölkerungsrepräsentative Daten existieren, die für Analysen zu den wirtschaftlichen und sozialen Verhältnissen von Selbstständigen geeignet sind, und die für diesen Zweck auch genutzt werden. Dennoch würde die vom BMAS geplante eigenständige Erhebung einen substanziellen Mehrwert schaffen. Erstens lässt sich damit ein im Vergleich zu den vorhandenen Datenquellen umfangreicheres, genauer auf die Erkenntnisinteressen des BMAS ausgerichtetes Fragenprogramm realisieren. In auf die Gesamtbevölkerung ausgerichteten Befragungen fehlen Items, die auf die besonderen Rahmenbedingungen und gesetzlichen Vorgaben für Selbstständige zugeschnitten sind. Deshalb lässt sich mit diesen nur unzureichend erfassen, wie gut diese Zielgruppe gegen Einkommens-, Alters- und Gesundheitsrisiken abgesichert ist. Zweitens können so die für belastbare differenzierte Analysen zu Teilpopulationen erforderlichen Fallzahlen realisiert werden.

Was den Feldzugang für eine bevölkerungsrepräsentative gesonderte Erhebung betrifft, muss die zu erreichende Stichprobe von der Personenebene und nicht von der Unternehmensebene ausgehend strukturiert werden. Dieselbe selbstständige Person kann mit mehreren Unternehmen gleichzeitig in Bezug stehen; andererseits stehen insbesondere Angehörige der Freien Berufe oft in keinem Bezug zu einem registrierten Unternehmen. Deshalb sollte eine bevölkerungsrepräsentative Zufallsstichprobe als Anker der geplanten Befragung gewählt werden. Damit mehr Gewerbetreibende mit einer größeren Zahl von Beschäftigten in die Nettostichprobe gelangen, könnten diese ergänzend anhand von umfassenden Unternehmensdatenbanken identifiziert und gezielt angesprochen werden. Bei beiden Ausgangsgesamtheiten ist allerdings in einem ersten Schritt ein Screening erforderlich, um die Bruttostichprobe mit teilnahmebereiten Zielpersonen für eine vertiefte Befragung zu gewinnen.

Der dabei zu betreibende Screening-Aufwand ist beträchtlich. Deswegen ist, wenn man von einer – mit Blick auf die Erfassung saisonaler Schwankungen zu empfehlenden – Feldzeit von einem Jahr und einer geringen Teilnahmebereitschaft in der Zielpopulation ausgeht, damit zu rechnen, dass sich nur eine Nettostichprobe mit rund 5.000 Fällen realisieren lässt. Diese Größenordnung erscheint aber ausreichend mit Blick auf das Ziel, differenzierte Aussagen zu den sozialen Lagen von relevanten Teilpopulationen selbstständig Erwerbstätiger zu gewinnen. Insgesamt ist bei Berücksichtigung eines angemessenen Vorlaufs zur Konzeption und Implementation der Befragung sowie zur analytischen Auswertung der gewonnenen Daten mit einem Zeitbedarf für die Durchführung der Studie von 22 Monaten ab Auftragsvergabe zu rechnen.

Um das multidimensionale Konzept der Prekarität und Faktoren, die prekäre Lagen von Selbstständigen begünstigen oder hemmen können, angemessen zu erfassen, wird ein umfangreiches Erhebungsprogramm benötigt. Es umfasst insbesondere die aktuelle Erwerbssituation von

Selbstständigen, Angaben zum selbstständig betriebenen Geschäft, zur bisherigen Erwerbsbiografie, zu Einkommen und vorhandenem Vermögen sowie zur sozialen oder privaten Vorsorge gegen die Risiken Alter, Einkommensverlust und Krankheit. Auch müssen differenzierte Informationen zum Haushaltskontext erhoben werden. Schließlich sollte das Fragenprogramm auch subjektive Einschätzungen, persönliche Einstellungen und Haltungen sowie Persönlichkeitsmerkmale aufnehmen – Größen, die in einem starken Zusammenhang mit der Inzidenz und Schwere prekärer sozialer Lagen stehen können. Das Erhebungsprogramm kann sich teilweise an Items orientieren, die in bevölkerungsrepräsentativen Umfragen gängig sind, so dass Vergleichsmöglichkeiten entstehen. Es muss aber auch eine Reihe spezifischer Items enthalten, insbesondere zu Besonderheiten der sozialversicherungsrechtlichen Absicherung von Selbstständigen, sowie zur von der Zielgruppe in der Regel zu betreibenden privaten Risikovorsorge.

Um diese umfangreichen und komplexen Befragungsinhalte angemessen umzusetzen, ist mit einer durchschnittlichen Interviewdauer von weit über einer halben Stunde zu rechnen. Damit scheidet selbstadministrierte Interviews in Papier- oder digitaler Form für die Kernerhebung aus. Vielmehr müssen telefonische oder persönliche computerassistierte Interviews geführt werden. Für das vorzuschaltende Screening einer bevölkerungsrepräsentativen Stichprobe bietet sich die Nutzung einer telefonischen Mehrthemenumfrage („Omnibus“) an. Werden zusätzlich Kontaktdaten aus einer Unternehmensdatenbank gewonnen, kommt für das Screening neben einem telefonischen Modus auch ein digitaler Modus in Betracht.

1. Einleitung

In Deutschland waren gemäß den Ergebnissen der Arbeitskräfteerhebung der Europäischen Union (AKE-EU) im Jahr 2020 insgesamt 3,5 Millionen Menschen selbstständig tätig. Darunter befanden sich rund 55 Prozent Solo-Selbstständige, das heißt Personen, die ihr eigenes Unternehmen führen oder ihren Beruf freischaffend selbstständig ausüben, ohne weitere Personen regulär zu beschäftigen. Dabei bilden Solo-Selbstständige eine heterogene Gruppe, die etwa Angehörige der traditionellen Freien Berufe und „neue“ Wissensarbeiter, aber auch wenig qualifizierte Anbieterinnen und Anbieter von einfachen Dienstleistungen umfasst. Für Solo-Selbstständigkeit ist charakteristisch, dass sie häufig auf persönlichen Wissensbeständen basiert und die Aufnahme eines Geschäfts gegen Gewinn nur vergleichsweise geringe personelle und wirtschaftliche Ressourcen erfordert (Schulze Buschoff et al. 2017). Dies kann unfreiwillige, prekäre solo-selbstständige Beschäftigung befördern (Westerveld 2012). Auch tragen Solo-Selbstständige im Vergleich zu anderen Selbstständigen in der Regel ein erhöhtes unternehmerisches Risiko.

Allgemein zeichnen sich Selbstständige, insbesondere Gründerinnen und Gründer, durch eine im Vergleich zu anderen Erwerbstätigen höhere Risikobereitschaft aus (Metzger 2021). Sie sind allerdings auch größeren wirtschaftlichen Risiken ausgesetzt als abhängig Beschäftigte. Sehr deutlich hat sich dies in der Corona-Krise gezeigt, von der selbstständig Erwerbstätige im Hinblick auf das Einkommen und die Chancen auf den Fortbestand des Geschäfts überdurchschnittlich betroffen waren (Seebauer et al. 2021; Schulze Buschoff und Emmeler 2021). Aber auch bei normaler Konjunkturlage streuen die Einkünfte aus selbstständiger Erwerbstätigkeit stärker als die Einkünfte aus abhängiger Beschäftigung. Dies betrifft insbesondere Solo-Selbstständige, von denen viele mit ihrem Geschäft nur ein geringes und unregelmäßiges Einkommen erzielen können (Bonin et al. 2022).

Selbstständig Erwerbstätige sind zudem nicht standardmäßig in die Sozialversicherungen integriert. Sie müssen also eigenständig Vorsorge für ihre Alterseinkünfte, den Krankheits- und Invaliditätsfall, sowie den Verlust von Erwerbseinkommen bzw. Zeiten der Auftragslosigkeit betreiben. Je nach den finanziellen Möglichkeiten, aber auch abhängig von der persönlichen Bereitschaft zur Vorsorge, kann die Absicherung zentraler wirtschaftlicher Risiken über private Versicherungen oder Vermögensbildung unzureichend bleiben, so dass Selbstständige mit wirtschaftlichen Schwierigkeiten eventuell lediglich über die Leistungen der sozialen Grundsicherung abgesichert sind.

Die sozialen (Problem-)Lagen von Selbstständigen sowie die sich daraus ergebenden besonderen Schutzbedarfe und die zu ihrer Erfüllung zielführenden sozialpolitischen Maßnahmen lassen sich für Deutschland bislang nur unzureichend statistisch fundiert beschreiben. Hierzu tragen verschiedene Faktoren bei. So ist der Anteil der selbstständig Erwerbstätigen an allen Erwerbstätigen relativ klein. Die Selbstständigenquote verharrt seit längerem bei Werten um zehn Prozent. Dadurch ergeben sich im Rahmen von auf die Gesamtbevölkerung ausgerichteten repräsentativen Erhebungen in der Regel zu geringe Fallzahlen, um zu belastbaren, sowohl nach Typen der selbstständigen Tätigkeit als auch nach sozio-demografischen Charakteristika differenzierten Ergebnissen zu kommen. Wegen kleiner Fallzahlen lässt sich insbesondere die Streuung bzw. Verteilung der von Selbstständigen erzielten Einkommen nur eingeschränkt analysieren. Weil der Anteil der Selbstständigen an der erwerbstätigen Bevölkerung eher klein ist, fehlen in auf die gesamte Bevölkerung ausgerichteten Befragungen zudem spezifische Items, die auf die besonderen Rahmenbedingungen und gesetzlichen Vorgaben für selbstständig Erwerbstätige zugeschnitten sind. Damit lässt sich bei dieser Zielgruppe insbesondere die Qualität der Absicherung gegen Einkommens-, Alters- und Gesundheitsrisiken bislang nur unvollständig erfassen.

Andere Erfassungsprobleme entstehen durch sogenannte „hybride Arbeitsverhältnisse“ (Bührmann et al. 2018). Dies sind Arbeitsverhältnisse, bei denen Erwerbstätige flexibel und in wechselnden Konstellationen abhängige und selbstständige Beschäftigungsverhältnisse miteinander verknüpfen. Beispielsweise begünstigt der verbreitete Projekt-Charakter der Arbeit im Bereich der Kunst- und Kreativwirtschaft ein „Multi-Jobbing“ (vgl. etwa Lingo und Tepper 2013), bei dem mehrere einzelne Tätigkeiten asynchron nebeneinander oder auch sequenziell nacheinander ausgeführt werden. Zur Hybridisierung kommt es, weil abhängige und selbstständige Beschäftigung unterschiedlich reguliert sind, etwa im Hinblick auf die Festlegung von Löhnen bzw. Entgelten oder den Zugang zu den sozialen Sicherungssystemen. Dass innerhalb des Bereichs der abhängigen Beschäftigung unterschiedliche Typen von Arbeitsverhältnissen existieren, macht die Hybridisierung noch komplexer.¹ Bei hybriden Konstellationen und stark diskontinuierlichen, komplexen Erwerbsverläufen („Portfolio-Karrieren“) wird die Inzidenz selbstständiger Erwerbsarbeit in den vorhandenen Daten tendenziell untererfasst. So zählt der Mikrozensus als selbstständig Erwerbstätige lediglich Personen, die entweder nur eine einzige Tätigkeit ausüben, die eine selbstständige Tätigkeit ist, oder deren Haupttätigkeit, definiert als die Tätigkeit mit der längsten Arbeitszeit, eine selbstständige Tätigkeit ist.

Vor diesem Hintergrund beabsichtigt das Bundesministerium für Arbeit und Soziales (BMAS), ein Forschungsprojekt durchzuführen mit dem Ziel, geeignete Daten zu erheben, auf deren Basis sich belastbare und differenzierte Einschätzungen zur sozialen Lage von Selbstständigen gewinnen lassen.² Dabei soll die geplante Untersuchung die Grundgesamtheit aller Personen in Deutschland abbilden, die in irgendeiner Form selbstständig tätig sind. Insbesondere sollen auch Personen, die nur im Nebenerwerb, geringfügig oder nur sporadisch selbstständig tätig sind, in die Erhebung einbezogen werden, um Einschätzungen zur Inzidenz und den Folgen hybrider Arbeitsverhältnisse zu ermöglichen. Die Festlegung auf die breitestmögliche Grundgesamtheit schließt eine gewisse Schwerpunktsetzung auf Formen der Selbstständigkeit, bei denen in besonderem Maße prekäre wirtschaftliche oder soziale Verhältnisse zu vermuten ist, nicht prinzipiell aus. Dazu wäre eine geschichtete Stichprobe zu realisieren. Diese Vorgehensweise setzt allerdings voraus, dass sich geeignete Faktoren zur Hochrechnung der Ergebnisse auf die Grundgesamtheit konstruieren lassen, die den systematisch unterschiedlichen Auswahlwahrscheinlichkeiten der Untersuchungseinheiten infolge der Schichtung Rechnung tragen.

¹ Zu differenzieren ist nach – befristeter oder unbefristeter – sozialversicherungspflichtiger Beschäftigung, geringfügig entlohnter Beschäftigung und – berufsmäßiger oder nicht berufsmäßiger – unständiger Beschäftigung, wenn Personen Beschäftigungen von weniger als einer Woche ausüben.

² Selbstverständlich müssen bei der geplanten Datenerhebung die Anforderungen des Datenschutzes, insbesondere die Vorgaben der Datenschutz-Grundverordnung (DSGVO) eingehalten werden. Dafür muss unter anderem gewährleistet sein, dass die gesamte Datenverarbeitung auf einem legitimen Zweck (hier dem berechtigten öffentlichen Interesse an den Forschungsergebnissen) beruht, die Daten nicht über den legitimen Zweck hinaus verarbeitet werden, und dass betroffene Personen über die Verarbeitungstätigkeiten bezüglich ihrer erfassten personenbezogenen Daten transparent informiert werden. Auch dürfen nur notwendige personenbezogene Daten erfasst werden, die zu löschen sind, sobald der legitime Zweck, für den sie erhoben wurden, erfüllt ist. Auf die Einhaltung der datenschutzrechtlichen Vorgaben muss das BMAS als Auftraggeber die mit der Durchführung des Forschungsvorhabens beauftragten Institutionen wie üblich verpflichten. Es ist davon auszugehen, dass die für die Erhebung und die wissenschaftliche Auswertung der Daten im vorliegenden Kontext fachlich geeigneten Unternehmen und Forschungseinrichtungen mit den Anforderungen des Datenschutzes sehr gut vertraut sind und deren Einhaltung gewährleisten können. Beispielsweise ist für einzuschaltende fachlich geeignete Befragungsunternehmen die Durchführung datenschutzkonformer bevölkerungsrepräsentativer Befragungen Standard; eventuelle Datenschutzverstöße würden das Geschäftsmodell elementar gefährden. Im Zuge eines eventuellen Vergabeverfahrens sollten die Bieter aufgefordert werden, ihre Maßnahmen und Konzepte zur Umsetzung der Anforderungen des Datenschutzes im spezifischen Kontext des Forschungsvorhabens dazulegen. Auch könnte sich das BMAS erforderliche Texte zur (mündlichen oder schriftlichen) Aufklärung der in die Untersuchung einbezogenen Personen im Verlauf der Durchführung der Untersuchung vor deren Nutzung zur Freigabe vorlegen lassen. Da keine besonderen Anforderungen oder Risiken bezüglich der bei der Durchführung des Forschungsvorhabens einzuhaltenden rechtlichen Rahmenbedingungen erkennbar sind, werden diese in der Machbarkeitsstudie nicht weiter betrachtet.

Inhaltlich soll der Schwerpunkt der geplanten Erhebung auf den potenziellen wirtschaftlichen und sozialen Problemlagen von selbstständig Tätigen liegen; dagegen sind die wirtschaftliche Lage und die Entwicklung der von Selbstständigen gegen Gewinn betriebenen Unternehmungen nur von sekundärem Interesse. Der zu entwickelnde Fragebogen soll insbesondere geeignet sein, prekäre wirtschaftliche und soziale Lebensverhältnisse zu erkennen. Folgt man der Definition von Schulze Buschoff et al. (2017, S. 57), ist „eine Beschäftigung [...] prekär, wenn sie mit einer erhöhten existenziellen Unsicherheit einhergeht“. Diese existenzielle Unsicherheit lässt sich anhand unterschiedlicher Risikofaktoren beurteilen. Hierzu zählen insbesondere eine geringe finanzielle Belastbarkeit in Folge eines relativ niedrigen Einkommens sowie eine unzureichende Absicherung gegen die zentralen sozialen Risiken Alter, Erwerbseinkommensausfälle, Krankheit und Invalidität. Existenzielle Unsicherheit kann sich aber auch in anderen Dimensionen, wie einer niedrigen Lebensqualität, schlechter Gesundheit oder hoher Unsicherheit über die Zukunft niederschlagen.³

Bei Prekarität handelt es sich somit um einen multidimensionalen Sachverhalt. Dennoch betrachten viele Studien zu Formen prekärer Beschäftigung einzelne Risikofaktoren – nicht zuletzt eine Folge nur begrenzt verfügbarer Daten. Jedoch liegt inzwischen auch eine Reihe von Untersuchungen vor, die der Multidimensionalität des Konzepts ausdrücklich Rechnung tragen und Indikatoren für Prekarität auf Grundlage einer Anzahl von Risikofaktoren konstruieren (vgl. etwa Rodgers und Rodgers 1989; Vosko et al. 2009; Standing 2011; Schulze Buschoff et al. 2017; Promberger et al. 2018).

Für die Messung von Prekarität bei Personen mit selbstständiger Erwerbstätigkeit ist es wichtig, die Ebenen des einzelnen Beschäftigungsverhältnisses, des Individuums und des Haushalts zu unterscheiden. Zum einen lässt sich aus der Beobachtung, dass eine Person eine selbstständige Beschäftigung mit einem relativ niedrigen Ertrag und fehlender Absicherung gegen wesentliche soziale Risiken ausübt, nicht ohne Weiteres schließen, dass ihre Lage innerhalb der Arbeitssphäre insgesamt prekär ist. Vielmehr müssen, um ein vollständiges Bild zu erhalten, auch die individuellen Einkünfte und Ansprüche auf Sozialleistungen erfasst werden, die sich im Zusammenhang mit weiteren Tätigkeiten in der Sphäre der abhängigen Beschäftigung ergeben können. Zum anderen muss der Kontext der Familie bzw. des eine Bedarfsgemeinschaft bildenden Haushalts berücksichtigt werden. Finanzielle und sonstige Sicherungsmechanismen können auch über die sozialrechtliche oder praktisch gelebte Einstandsgemeinschaft greifen. Ebenfalls können individuell gesehen nicht prekär Beschäftigte außerhalb ihrer Arbeitssphäre über den Haushaltszusammenhang mit prekären Lebenslagen konfrontiert sein.

³ Die zuletzt genannten Dimensionen erfassen mögliche Folgen wirtschaftlich und sozial schlecht abgesicherter Beschäftigung; jedoch kann es sich zugleich auch um Faktoren handeln, die die Ausübung solcher Beschäftigung begünstigen. Ursache und Wirkung müssen mittels geeigneter empirischer Methoden herausgearbeitet werden, etwa durch die Bewertung von Veränderungen in den Zieldimensionen bei Veränderungen der Form der Beschäftigung.

Die folgenden Tabellen zeigen, wie prekäre Lagen von selbstständig Tätigen konkret multidimensional gemessen werden können.⁴ Dabei umfasst Tabelle 1.1 Dimensionen der Prekarität der individuellen Beschäftigungslage; hier muss bei der Diagnose gegebenenfalls über mehrere (parallel ausgeübte oder zeitlich eng aufeinander folgende) Beschäftigungsverhältnisse hinweg aggregiert werden. Tabelle 1.2 konkretisiert Dimensionen prekärer Haushaltslagen, mit denen selbstständig Tätige konfrontiert sein können, selbst wenn ihre eigene Position in der Arbeitswelt nicht als prekär zu klassifizieren ist. Für jeden der in beiden Tabellen enthaltenen Indikatoren ist eine Grenze zu ziehen, wann ein Zustand als prekär oder nicht-prekär einzustufen ist. Die möglichen Abgrenzungskriterien sind letztlich ad hoc und sollten deswegen jeweils in plausiblen Grenzen variiert werden, um die Sensitivität der empirischen Befunde zu prüfen.

Tabelle 1.1 Indikatoren zur Messung der Dimensionen von prekärer Beschäftigung auf der Personenebene

Dimensionen prekärer Beschäftigung	Indikator	Mögliche Abgrenzung prekärer Lage
Einkommen	Niedriglohn	Bruttoentgelt je Stunde < zwei Drittel Median der Bruttostunden-Entgelte
	Existenzsicherndes Einkommen	Individuelles Bruttojahreseinkommen aus Beschäftigung < Grundfreibetrag Einkommensteuer
Mangelnde soziale Absicherung	Absicherung gegen Einkommensrisiken	Beschäftigung ohne eigene Beiträge zur ALV, ohne Versicherung gegen Invalidität / Berufsunfähigkeit
	Absicherung gegen Altersrisiken	Beschäftigung ohne eigene Beiträge zur GKV / regelmäßige private Altersvorsorge
	Absicherung gegen Gesundheitsrisiken	Beschäftigung ohne eigene Beiträge zur GKV / PKV
Arbeitsqualität	Einfache Arbeit	Beschäftigung erfordert keine berufliche Qualifikation / unterhalb der eigenen beruflichen Qualifikation
	Unsicherheit	Hohes Risiko des Verlusts der Beschäftigung / des Arbeitsplatzes; Hohe Volatilität des Einkommens bei niedrigem Durchschnitt
	Gesundheitsrisiko	Ausgeübter Beruf mit hoher Exposition gegenüber Belastungen gemäß Job-Exposure Matrix

Quelle: Eigene Darstellung des IZA in Anlehnung an Promberger et al. (2018: 12).

⁴ Der Ausgangspunkt der Darstellung ist das multidimensionale Messkonzept, das von Promberger et al. (2018) für eine Analyse der Verfestigung von Prekarität bei Erwerbstätigen in Deutschland entwickelt wurde.

Tabelle 1.2 Indikatoren zur Messung der Dimensionen von prekären Lagen auf der Haushaltsebene

Dimensionen prekärer Lage	Indikator	Mögliche Abgrenzung prekärer Lage
Finanzielle Lage	Armutrisiko	Bedarfsgewichtetes Einkommen unter 60% des Medians
	Finanzielle Rücklagen	Haushalt spart nicht regelmäßig / kann nur kurze Zeiträume ohne Arbeit/Aufträge mit eigenen finanziellen Mitteln überbrücken
	Schulden	Hoher Anteil des bedarfsgewichteten Einkommens wird für Zins- und Tilgungszahlungen aufgewendet
Indirekte soziale Sicherung	Abgeleitete Ansprüche	Sozialversicherungspflichtige Beschäftigung des Ehepartners / der Ehepartnerin
Wohnen	Beengte Wohnverhältnisse	Niedrige Wohnfläche pro Person im Haushalt
	Eigentum	Haushalt besitzt selbst- oder fremdgenutzte Immobilie
	Miete	Hoher Anteil der (Warm-)Miete am bedarfsgewichteten Einkommen
Mobilität	Zugang zu Verkehrsmitteln	Haushalt verfügt über Auto

Quelle: Eigene Darstellung des IZA in Anlehnung an Promberger et al. (2018: 12).

Weil Prekarität dadurch entsteht, dass verschiedene einzelne Risikofaktoren zugleich auftreten (Rodgers und Rodgers 1989), ergibt sich darüber hinaus ein Aggregationsproblem. Eine einfache Lösung wäre die Bildung eines Summenindex, möglichst getrennt für Risikofaktoren auf der individuellen und der Haushaltsebene, bei dem die einzelnen Teilindikatoren gleich stark gewichtet werden. Konzeptionell vorzuziehen sind allerdings komplexere Aggregationsverfahren, die den Korrelationen der verschiedenen Risikofaktoren Rechnung tragen, wie etwa Cluster-Analysen (Schulze Buschoff et al. 2017).

Im Idealfall sollten Risikofaktoren, die auf individueller und Haushaltsebene zu prekären Lagen beitragen können, im Zeitverlauf betrachtet werden. So lässt sich beurteilen, inwieweit es sich bei beobachteten prekären Lagen um einen im (erwerbs-)biografischen Verlauf nur kurz andauernden, wiederkehrenden, sich verfestigenden oder dauerhaften Zustand handelt. Per Längsschnittanalyse lassen sich zudem präziser die Faktoren bestimmen, die das Auftreten und die Verfestigung prekärer Lagen auf der Mikroebene der Individuen und Haushalte begünstigen oder hemmen.

Falls die Erkenntnisinteressen des BMAS im Hinblick auf die prekären Lagen von selbstständig tätigen Personen nicht bereits auf Grundlage vorhandener Panel-Daten bedient werden können, kommt die Neuanlage einer spezifischen Wiederholungsbefragung von Individuen oder Haushalten wohl dennoch nicht in Frage. Der zeitliche Vorlauf, bis Auswertungen im Längsschnitt vorgenommen werden könnten, wäre zu groß, und die Erhebungskosten würden sich naturgemäß vervielfachen.

Wenn zur Erfüllung der Erkenntnisinteressen des BMAS eine spezifische Querschnittserhebung erforderlich wird, sollte diese zumindest so angelegt werden, dass die Untersuchungseinheiten gegebenenfalls für die Zwecke einer eventuellen Anschlussstudie, die Dynamiken im Zeitverlauf fokussiert, erneut kontaktiert werden können.

Diese Expertise zeigt im Sinne einer Machbarkeitsstudie, wie ein zielführendes und wirtschaftliches Erhebungsdesign für das geplante Forschungsvorhaben des BMAS aussehen und zeitnah umgesetzt werden könnte. Im Zentrum stehen dabei Aspekte des Zugangs zur Grundgesamtheit, des Erhebungsmodus sowie der Fragebogengestaltung. Zudem werden Überlegungen zu den für die Umsetzung des Vorhabens erforderlichen zeitlichen und finanziellen Ressourcen angestellt. Für die Expertise wurden, ausgehend von einer per Desk Research erfolgten Sichtung und Bewertung bereits vorliegender Studien und Datensätze zur Beurteilung der sozialen Lage von Selbstständigen in Deutschland, eine Reihe von Expertengesprächen geführt. Darüber hinaus wurden Gestaltungsoptionen und bei der Umsetzung zu berücksichtigende praktische Restriktionen mit fachlich einschlägig erfahrenen Wissenschaftlerinnen und Wissenschaftlern sowie Umfrage-Praktikerinnen und -Praktikern in einem Design-Workshop am 12. April 2022 unter Beteiligung des BMAS diskutiert.⁵

Die weitere Expertise gliedert sich wie folgt: Kapitel 2 fasst zusammen, welche Datenquellen und Studien mit Bezug zur sozialen Lage von Selbstständigen in Deutschland bereits vorhanden sind, die als Anknüpfungspunkt für eine weiterführende Untersuchung dienen könnten. Kapitel 3 legt dar, welche Möglichkeiten für die Anlage einer neuen Erhebung bei Selbstständigen bestehen, und erörtert deren jeweilige Vor- und Nachteile. Kapitel 4 skizziert die Befragungsinhalte, die für eine Untersuchung zur sozialen Lage von Selbstständigen abgedeckt werden sollten. Schließlich diskutiert das abschließende Kapitel 5 zusammenfassend die Durchführung des Vorhabens auf Grundlage der Überlegungen der vorangegangenen Kapitel.

⁵ An diesem Design-Workshop waren neben dem IZA beteiligt: Deutsches Institut für Wirtschaftsforschung Berlin (DIW Berlin), forsa Gesellschaft für Sozialforschung und statistische Analysen Berlin, ifas Institut für angewandte Sozialwissenschaft Bonn, Institut für Arbeitsmarkt und Berufsforschung Nürnberg (IAB), Kantar Public Public Division München, SOKO Institut für Sozialforschung & Kommunikation Bielefeld, Leibniz-Zentrum für Europäische Wirtschaftsforschung (ZEW).

2. Datenquellen und Studien mit Bezug zur sozialen Lage von Selbstständigen

Zur Erfassung der Ausgangslage werden in diesem Abschnitt vorhandene Datenquellen zur wirtschaftlichen und sozialen Lage von Selbstständigen in Deutschland beschrieben und deren Qualitätsmerkmale – insbesondere im Hinblick auf Repräsentativität und Fallzahlen, Querschnitt- und Längsschnitt-Dimension sowie Merkmalspektrum – diskutiert. Dabei werden sowohl etablierte laufende und abgeschlossene Befragungen als auch neu durchgeführte Umfragen, die meist in Zusammenhang mit der Corona-Pandemie entstanden sind, berücksichtigt.

2.1 Sozio-ökonomisches Panel

Die am häufigsten verwendete Datenquelle für Auswertungen zur sozialen Lage, zur sozialen Absicherung oder zum Vorsorgeverhalten von Selbstständigen in Deutschland ist das Sozio-ökonomische Panel (SOEP). Das SOEP ist eine Panelstudie, für die jährlich etwa 30.000 Menschen in knapp 15.000 Haushalten befragt werden. Bonin et al. (2022) nutzen – neben weiteren Datenquellen – das SOEP, um insbesondere die Einkommenssituation und die Vermögensbildung von Selbstständigen in Deutschland zu analysieren. Hierbei zeigen Daten aus dem Jahr 2020, dass Solo-Selbstständigkeit neben einer hohen Teilzeitquote auch durch ein relativ geringes Einkommen gekennzeichnet ist. Berücksichtigt man den Haushaltskontext, ist das Einkommen aus solo-selbstständiger Tätigkeit allerdings nur ein Nebenverdienst, der ein höheres Einkommen der Partnerin oder des Partners ergänzt. Gemessen am bedarfsgewichteten Nettoäquivalenzeinkommen⁶ ist das verfügbare Haushaltseinkommen von Solo-Selbstständigen ähnlich hoch wie bei abhängig Beschäftigten – am unteren Rand der Einkommensverteilung sogar etwas höher. Was das Vermögen betrifft, befinden sich Solo-Selbstständige in der mittleren Position zwischen Selbstständigen mit Beschäftigten und abhängig Beschäftigten.

Frühere Studien zur Einkommens- und Vermögenssituation von (Solo-)Selbstständigen, die Daten aus dem Jahr 2010 und 2012 nutzen, sind etwa Niehues und Pimpertz (2012) und Beznoska (2017). So finden Niehues und Pimpertz (2012), dass viele Solo-Selbstständige in Teilzeit erwerbstätig sind und nur als Zweitverdiener zum Haushaltseinkommen beitragen. Obwohl sie mit ihrer Tätigkeit nur ein relativ geringes eigenes Einkommen erzielen, sind sie – gemessen am äquivalenzgewichteten Einkommen des gesamten Haushalts – unterdurchschnittlich oft armutsgefährdet. Beznoska (2017) nutzt Daten eines Erhebungsschwerpunkts zum Thema Vermögen im SOEP aus dem Jahr 2012 und vergleicht die Vermögensverteilung der Selbstständigen mit derjenigen von sozialversicherungspflichtig Beschäftigten. Hierbei wird nicht zwischen Solo-Selbstständigen und Selbstständigen mit Beschäftigten unterschieden. Die Ergebnisse der Studie zeigen, dass das Median-Nettovermögen in allen Altersgruppen deutlich über demjenigen der sozialversicherungspflichtig Beschäftigten liegt. Auch Bonin et al. (2022) zeigen, dass insbesondere im Hinblick auf den Immobilienbesitz als Vermögensbestandteil Selbstständige mit Beschäftigten bessergestellt sind als sozialversicherungspflichtige Beschäftigte.

Verschiedene Studien haben außerdem die Auswirkungen der Corona-Pandemie auf die Gruppe der Selbstständigen ausgewertet. Kritikos et al. (2020) etwa nutzen Daten der SOEP-CoV-Studie, für die

⁶ Das Nettoäquivalenzeinkommen korrigiert Struktureffekte der Haushaltszusammensetzung, indem das beobachtete Nettoeinkommen je nach Anzahl und Alter der Haushaltsmitglieder gewichtet wird. Die verwendeten Gewichte reflektieren den unterschiedlichen Bedarf der verschiedenen Personen im Haushalt. Die Bedarfsgewichtung gewährleistet, dass verschieden zusammengesetzte Haushalte mit gleichem Nettoäquivalenzeinkommen näherungsweise über die gleiche wirtschaftliche Leistungsfähigkeit verfügen.

eine Teilstichprobe des SOEP im Frühjahr 2020 sowie im Januar und Februar 2021 zu spezifischen Themen rund um die Corona-Pandemie befragt wurde. Die Ergebnisse zeigen, dass rund 60 Prozent der Selbstständigen von Einkommensverlusten berichten, während dies nur etwa 15 Prozent der abhängig Beschäftigten tun. Zudem berichten Selbstständige deutlich häufiger von großen Sorgen um die eigene wirtschaftliche Situation als abhängig Beschäftigte. Seebauer et al. (2021) zeigen darüber hinaus, dass weibliche Selbstständige noch einmal stärker von den Auswirkungen der Corona-Pandemie betroffen sind als männlich Selbstständige, weil sie überproportional häufig in den von der Corona-Pandemie besonders betroffenen Branchen tätig sind.

Die SOEP-Daten bieten bei der Analyse der sozialen Lagen von Selbstständigen mehrere Vorteile. So ist die Datenbasis repräsentativ, und die Panel-Dimension ermöglicht Längsschnittanalysen, also Vergleiche derselben Personen über die Zeit. Dadurch können etwa Wege in die und aus der Selbstständigkeit analysiert werden. Es existiert ein breites Fragenspektrum, das insbesondere auch verschiedene Dimensionen von prekärer Beschäftigung und prekärer Lagen auf Haushaltsebene umfasst. Zudem bestehen Vergleichsmöglichkeiten mit anderen Bevölkerungsgruppen, wie etwa mit der Gruppe der abhängig Beschäftigten. Solo-Selbstständige und Selbstständige mit Beschäftigten können unterschieden werden, und es besteht eine recht große Flexibilität hinsichtlich der Definition des Status „selbstständig“. So werden beispielsweise hybride Erwerbskonstellationen recht präzise erfasst, da neben Angaben zur beruflichen Stellung in der Haupttätigkeit auch diesbezügliche Angaben für bis zu drei Nebentätigkeiten erhoben werden. Schließlich ist das SOEP eine etablierte Datenbasis, die regelmäßig für die Analysen sozialer Lagen verwendet wird, etwa im Rahmen des Sechsten Armuts- und Reichtumsberichts (Bundesregierung 2021).

Allerdings hat das SOEP mit Bezug auf die Fragestellungen der geplanten Studie auch bedeutsame Schwächen. Der größte Nachteil ist eine unzureichende Fallzahl. Im aktuellen Querschnitt werden lediglich etwa 1.000 selbstständig erwerbstätige Personen erfasst. Damit werden bei disaggregierten Auswertungen und spezifischen Analysen für kleinere Teilpopulationen von Selbstständigen schnell Grenzen erreicht. Beispielsweise sind belastbare Aussagen im Hinblick auf die Teilpopulation der hybrid Selbstständigen, die nur einen Anteil von etwa zwei Prozent an allen Erwerbstätigen ausmacht, kaum möglich. Auch Unternehmerinnen und Unternehmer mit einer größeren Anzahl von Beschäftigten sind – gemäß ihrem Anteil an der Gesamtbevölkerung – im SOEP nur relativ selten anzutreffen. Ein anderer Nachteil ist, dass die regulären Fragebögen, wie angesichts des niedrigen Anteils an selbstständig Erwerbstätigen in der Stichprobe zu erwarten, nur wenige Fragen beinhalten, die sich speziell auf Aspekte der Selbstständigkeit beziehen. Insbesondere werden die Dimensionen der sozialen und privaten Absicherung gegen zentrale Risiken (Alter, Einkommensverlust, Krankheit), bei denen Selbstständige mit speziellen institutionellen und rechtlichen Rahmenbedingungen konfrontiert sind, nicht so differenziert erfasst, wie es für die Zwecke einer spezifischen Untersuchung zur sozialen Lage von selbstständig Erwerbstätigen erforderlich ist (vgl. Abschnitt 4).

2.2 Panel Arbeitsmarkt und soziale Sicherung

Das vom IAB durchgeführte Panel Arbeitsmarkt und soziale Sicherung (PASS) dient primär dem Ziel, die Dynamik des Grundsicherungsbezugs und die soziale Lage von Haushalten im Grundsicherungsbezug zu untersuchen. Das PASS umfasst neben einer jährlich um Neuzugänge zur Grundsicherung aufgefrischten Teilstichprobe von Bedarfsgemeinschaften auch eine Teilstichprobe in Form einer allgemeinen Bevölkerungstichprobe, mit der eine Hochrechnung der PASS-Befunde auf die Wohnbevölkerung in Deutschland möglich wird.

Der Ausrichtung des PASS entsprechend, lässt sich damit insbesondere die Situation von Selbstständigen untersuchen, die zusätzlich zu ihrem Einkommen auf staatliche Leistungen angewiesen sind. So nutzen Pahnke et al. (2019) diesen Datensatz, um zu zeigen, dass erlittene

Einkommensverluste aus der selbstständigen Tätigkeit häufig die Ursache für den Eintritt in die Grundsicherung sind, wobei eine vorhandene Partnerschaft und eine gute Ausbildung das Risiko dieses Eintritts verringern. Die Einkommen von selbstständigen Leistungsbeziehenden sind im Vergleich zu denjenigen von abhängig beschäftigten Leistungsbeziehenden deutlich geringer. Die meisten Selbstständigen verlassen die Grundsicherung allerdings auch wieder recht zügig innerhalb der ersten beiden Jahre des Leistungsbezugs.

Mit Blick auf die Erkenntnisziele der vom BMAS geplanten Studie hat das PASS den Vorteil, dass sich damit die Lebensumstände von armutsgefährdeten und auf Leistungen des Grundsicherungssystems angewiesenen Personen gut untersuchen lassen. Es ist von der Anlage der Untersuchung davon auszugehen, dass prekär Erwerbstätige, und damit auch prekär Selbstständige, in den Daten relativ stärker vertreten sind als im SOEP. Wie im SOEP bietet die Panel-Dimension zusätzlich den Vorteil, dass auch dynamische Verläufe in die (prekäre) Selbstständigkeit oder daraus hinaus betrachtet werden können.

Im Hinblick auf den Zweck der geplanten Untersuchung hat das PASS jedoch auch gravierende Nachteile. So ist die realisierte Panelstichprobe mit zuletzt etwa 10.400 Personen und 7.800 Haushalten relativ klein – die Anzahl der erfassten Selbstständigen entsprechend noch einmal kleiner als im SOEP. Zum Fallzahlenproblem kommt hinzu, dass der Anlage der Stichprobe entsprechend Selbstständige außerhalb der Grundsicherung unterproportional vertreten sind. Es gibt damit nur relativ wenige Beobachtungen, um die Verhältnisse bei wirtschaftlich und sozial besser gestellten Selbstständigen hochzurechnen. Dies läuft dem Ziel entgegen, ein aussagekräftiges vollständiges Bild der sozialen Lagen von Selbstständigen in Deutschland zu zeichnen. Zudem fehlen, wie im SOEP, im regulären Fragenprogramm wesentliche Items zur spezifischen Bewertung der sozialen Absicherung von Selbstständigen gegen zentrale Lebensrisiken.

2.3 Mikrozensus

Der Mikrozensus ist eine repräsentative Haushaltsbefragung der amtlichen Statistik in Deutschland, in der rund ein Prozent der Bevölkerung zu ihren Arbeits- und Lebensbedingungen befragt wird. Statistische Auswertungen auf Basis des Mikrozensus geben daher Auskunft über die Situation aller Erwerbstätigen und erlauben so einen Vergleich zwischen Selbstständigen und abhängig Beschäftigten.

Bonin et al. (2022) nutzen den Mikrozensus vor allem, um die Entwicklung der selbstständigen Beschäftigung allgemein sowie ihre Aufteilung nach Branchen und sozio-demografischen Merkmalen darzustellen. Die Ergebnisse machen deutlich, dass sich die Gruppe der Selbstständigen seit dem Jahr 2002 mehrheitlich aus Solo-Selbstständigen zusammensetzt. Der Anteil der Selbstständigen an allen Erwerbstätigen lag in den vergangenen 20 Jahren relativ konstant zwischen knapp zehn und elf Prozent. Der Stellenwert von selbstständiger Erwerbstätigkeit ist jedoch je nach sozio-demografischer Gruppe verschieden hoch und verändert sich teilweise auch über die Zeit. So sind Frauen weniger häufig selbstständig tätig als Männer, während die Selbstständigenquote von Personen mit deutscher Staatsangehörigkeit in etwa derjenigen von Personen mit ausländischer Staatsangehörigkeit entspricht. Zudem sind ältere Erwerbstätige häufiger selbstständig tätig als jüngere Erwerbstätige.

Rosemann und Koch (2012) verwenden Daten des Mikrozensus, um die sozio-ökonomische Situation der Solo-Selbstständigen in der Kreativwirtschaft zu untersuchen.

Die Vorteile des Mikrozensus liegen in der großen repräsentativen Stichprobe, die Analysen verschiedener Teil- und Untergruppen ermöglicht, wie etwa Analysen nach Branchen und sozio-

demografischen Merkmalen sowie regional differenzierte Analysen. Die bei den zuvor behandelten Panel-Studien limitierenden Fallzahlenprobleme stellen sich hier nicht. Darüber hinaus wird der aktuelle Erwerbsstatus (inklusive Selbstständigkeit) im Erhebungsteil zur Arbeitsmarktbeteiligung sehr differenziert erfasst. Ebenfalls gibt es einen Erhebungsteil zu den Themen Einkommen und Lebensbedingungen, in dessen Rahmen unter anderem Angaben zu materiellen Entbehrungen, sozialer Teilhabe und zur Gesundheit erhoben werden, anhand derer sich Einschätzungen zu Armutsrisiken und sozialer Ausgrenzung gewinnen lassen.

Mit Blick auf die Ziele der geplanten Studie zur sozialen Lage von Selbstständigen weist jedoch auch der Mikrozensus erhebliche Nachteile auf. Erstens liegen Einkommensinformationen nur in kategorialer und damit zu grober Form vor. Zweitens werden nur wenige Angaben zum Haushalt erhoben; insbesondere fehlen wichtige Informationen zu den Erwerbsverhältnissen und Einkommen von Angehörigen. Drittens werden Vermögen nur rudimentär erfasst. Viertens liegen nur unvollständige Angaben zur sozialen Absicherung vor – Fragen zur Krankenversicherung etwa werden nur alle vier Jahre erhoben. Schließlich ist der Mikrozensus primär als Querschnitterhebung mit Fokus auf die Verhältnisse zum Zeitpunkt der Befragung konzipiert, so dass sich erwerbsbiografische Verläufe und wirtschaftliche Dynamiken auf der Personenebene damit nur stark eingeschränkt beurteilen lassen.

2.4 HBS-Erwerbstätigenbefragung

Die Hans-Böckler-Stiftung (HBS) hat während der Corona-Krise begonnen, die Arbeits- und Lebensbedingungen von Erwerbspersonen in Deutschland regelmäßig zu erfassen. Bisher gab es sieben (Panel-)Wellen der vor diesem Hintergrund initiierten Erwerbstätigenbefragung, zuletzt im Januar 2022. Die Interviews werden als computergestützte Online-Interviews (CAWI) mit Erwerbspersonen ab 16 Jahren durchgeführt. Schulze Buschoff und Emmeler (2021) nutzen diese Daten, um die Auswirkungen der Corona-Krise auf Selbstständige in Deutschland zu untersuchen. Ähnlich wie Kritikros et al. (2020) ermittelt auch diese Studie, dass Selbstständige ihre Arbeitssituation und ihre finanzielle Situation häufiger als stark belastend empfinden als abhängig Beschäftigte und auch häufiger von Einkommensverlusten berichten.

Ausgangspunkt dieser Befragung war die Ziehung einer Quotenstichprobe im Rahmen eines Online-Access-Panels, um die Erwerbspersonen in Deutschland basierend auf Sollzahlen der amtlichen Statistik abbilden zu können. Die Stichprobe stellt dennoch keine Zufallsstichprobe dar, da die Auswahlgesamtheit nicht repräsentativ ist: Zur Auswahlgesamtheit gehören nur Personen, die eine Payback-Karte besitzen. Um das Risiko des Selektionsproblems bei Online-Befragungen zu reduzieren, wurde jedoch auch offline rekrutiert. Selbstständige waren in den ersten vier Wellen der HBS-Erwerbspersonenbefragung jedoch stark unterrepräsentiert. Lediglich 208 Befragte und damit etwas mehr als vier Prozent des (erwerbstätigen) Samples aus Welle 5, welches bereits in Welle 1 befragt wurde, wurden als Selbstständige identifiziert. Nach Bonin et al. (2022) liegt dieser Anteil jedoch eher bei rund zehn Prozent. Für die fünfte Erhebungswelle wurde daher eine disproportionale Nachziehung von Selbstständigen durchgeführt, wodurch im Juli 2021 Angaben zu insgesamt 1.350 Selbstständigen vorlagen.

Die HBS-Erwerbstätigenbefragung hat den Vorteil einer großen Aktualität und einer hohen Frequenz. Allerdings erscheint fraglich, ob die Erhebung nach dem Abflauen der Corona-Pandemie noch weitergeführt wird. Die Befragung zeigt darüber hinaus die gravierenden Probleme, die generell bei einer Rekrutierung von Erwerbstätigen im Rahmen eines Online-Access-Panels zu erwarten sind. Obwohl Quotenvorgaben getroffen wurden, um die Bevölkerung nach den Merkmalen Alter, Geschlecht, Bundesland und Bildung abbilden und die Ergebnisse auf die Sollzahlen der amtlichen Statistik hochrechnen zu können, wurde –wie sich an der zunächst starken Untererfassung der

Selbstständigen zeigt – keine adäquat repräsentative Stichprobe erreicht. Inwieweit die nachgezogene Stichprobe von Selbstständigen mit Selektivitätsproblemen behaftet ist, lässt sich anhand der vorliegenden Methodenberichte und Datenreports nicht beurteilen. Die aufgetretenen Probleme bei der Rekrutierung weisen zudem auf die eine erhöhte Zögerlichkeit bei Selbstständigen hin, an Befragungen teilzunehmen. Die muss bei der Stichprobenplanung grundsätzlich beachtet werden.

2.5 IAB-Studien

Das IAB hat wiederholt die freiwillige Inanspruchnahme der Arbeitslosenversicherung durch Selbstständige untersucht. Hierfür wurde die Möglichkeit genutzt, die Integrierten Erwerbsbiografien (IEB), die tagesgenaue Informationen zu Leistungsbezug, sozialversicherungspflichtiger Beschäftigung und Maßnahmenteilnahmen enthalten, mit Daten von Selbstständigen zum Versicherungspflichtverhältnis auf Antrag aus den operativen Daten der BA zu verknüpfen (Oberfichtner 2019; Jahn und Springer 2013). So wird etwa die Arbeitslosenversicherung für Existenzgründer untersucht (Oberfichtner 2019) und die Ergebnisse zeigen, dass die Versicherungsneigung mit dem Alter, der Dauer des Arbeitslosengeldbezugs im Jahr vor der Gründung, dem Bildungsstand und dem früheren Lohn zunimmt. Diese IAB-Daten enthalten jedoch keine direkten Angaben zur selbstständigen Tätigkeit, wie etwa den Zeitpunkt des ersten getätigten Umsatzes. Es wird damit nur eine recht spezielle Gruppe von Selbstständigen erfasst.

Das IAB hat außerdem eine Online-Befragung unter Gründerinnen und Gründern durchgeführt, die von April bis Juli 2017 in den Gründungszuschuss zugegangen sind (Jahn und Springer 2013). Etwa einen Monat vor dem Ende der ersten Förderphase wurde danach gefragt, ob sie sich freiwillig arbeitslosenversichert haben, was auf 22 Prozent dieser Personen zutraf. Die Komplexität der Regelungen und die fehlende Planbarkeit, etwa über die konkrete Versicherungsleistung, stellen häufig genannte Gründe dar, sich nicht zu versichern. Auch bei dieser Befragung wird nur eine spezielle Gruppe von Selbstständigen erfasst.

Diese Beispiele zeigen, dass Selbstständige auch in den Statistiken der BA, die für wissenschaftliche Zwecke zugänglich sind, gewisse Spuren hinterlassen. Allerdings geschieht dies bei selbstständig Beschäftigten nur selten, weil diese in der Regel nicht in die Sozialversicherungen einbezogen sind und somit auch keine Meldungen zu Beschäftigungen, Leistungsempfang, Maßnahmenteilnahmen oder Arbeitssuche anfallen, die die Basis für die IEB bilden. Während die Datenbestände der BA in der Arbeitsmarktforschung als Zugangsweg für Befragungen von abhängig Beschäftigten häufig genutzt werden, kommt dieser Ansatz wegen der, wenn überhaupt, extrem selektiven Erfassung von Selbstständigen in den BA-Daten für die Zwecke der geplanten Untersuchung nicht in Frage. Allenfalls lässt sich, wie in Abschnitt 5.8 noch näher ausgeführt wird, im Anschluss an die durchzuführende Erhebung eine Verknüpfung zu den IEB-Daten herstellen, um genauere Daten über vergangene Episoden abhängiger Beschäftigung (und damit auch Versicherungszeiten in der Renten- und Arbeitslosenversicherung) und eine eventuelle Teilnahme an aktiven arbeitsmarktpolitischen Maßnahmen der befragten Selbstständigen zu erhalten.

2.6 Spezielle Befragungen von Selbstständigen

Im April und Mai 2020 wurde vom Verband der Gründer und Selbstständigen Deutschland (VGSD), dem ZEW und dem Forschungszentrum Mittelstand an der Universität Trier eine Online-Umfrage von mehr als 27.000 Selbstständigen durchgeführt. Vom VGSD wurden personalisierte E-Mails an Mitglieder versandt, und es wurde im Newsletter sowie über andere Verbände für die Umfrage geworben. Das Auswahlverfahren der Umfrage entspricht demnach einem Schneeball-Verfahren, um möglichst viele Selbstständige zur Teilnahme akquirieren zu können. Die Stichprobe

ist groß genug, um spezifische Auswertungen für Solo-Selbstständige vorzunehmen, wie etwa in Bertschek und Erdsiek (2020), die die Auswirkungen der Corona-Krise auf Selbstständige untersuchen. Ihre Studie wertet Antworten von 16.000 hauptberuflich Solo-Selbstständigen im Alter zwischen 24 und 65 Jahren aus. Die Ergebnisse zeigen, dass insbesondere in konsumnahen Branchen, wie Gastronomie und Beherbergung sowie Events und Veranstaltungen, etwa 90 Prozent der Solo-Selbstständigen von erheblichen Umsatzeinbußen von mehr als 75 Prozent berichten und demnach sehr stark von der Corona-Krise betroffen sind. Ein hoher Digitalisierungsgrad des Angebots scheint mit einer stärkeren Krisenresistenz verbunden zu sein.

Eine weitere Befragung von Selbstständigen, die während der Corona-Pandemie initiiert wurde, ist eine Pop-up-Befragung von KfW Research und der Gründerplattform. Die seit 2018 betrieben und vom Bundesministerium für Wirtschaft und Klimaschutz (BMWK) und der KfW geförderte Gründerplattform, auf der etwa Geschäftsmodell und Businessplan entwickelt und getestet werden können, wird meist von jungen Selbstständigen genutzt, die erst seit wenigen Jahren selbstständig tätig sind. Bisher gab es zwei Befragungswellen. Im Jahr 2020 (2021) haben 429 (492) Selbstständige und 167 (218) Gründungsinteressierte teilgenommen. Für den Zeitraum der „Blitzbefragungen“ wird jeweils ein Befragungslink auf der Gründerplattform freigeschaltet. Die Daten sind daher nicht repräsentativ, sollen laut KfW Research jedoch einen wichtigen Einblick in die Auswirkungen der Corona-Krise auf Selbstständige und Gründungsinteressierte geben können (Metzger 2020, 2021).

Im Rahmen eines Forschungsprojekts mit dem Titel „Self-Employment without Personnel: Between Freedom and Insecurity“, das in Kooperation mit dem Wirtschafts- und Sozialwissenschaftlichen Institut (WSI) in der Hans-Böckler-Stiftung an der Universität Utrecht durchgeführt wurde, fand eine Befragung von Solo-Selbstständigen in Form einer computergestützten Web-Befragung (CAWI) statt. Dabei wurden 757 Interviews in Deutschland und 793 Interviews in den Niederlanden geführt, die ländervergleichende Analysen ermöglichen. Auf Grundlage dieser Daten analysieren Schulze Buschoff et al. (2017), welche Faktoren die Prekarität von solo-selbstständig Tätigen erklären können, wobei mittels eines Cluster-Verfahrens die Dimensionen einer prekären finanziellen Situation und eines Mangels an sozialer Absicherung unterschieden werden. Im Ergebnis zeigt sich, dass Solo-Selbstständigkeit in den Niederlanden seltener prekär, häufiger Gewinn erwirtschaftend und besser abgesichert ist als in Deutschland.

Das erste Beispiel zeigt, dass es durchaus möglich ist, mittels CAWI-Verfahren eine große Anzahl von Selbstständigen für eine Befragung zu erreichen. Allerdings ist der Rekrutierungsprozess hier, wie auch im zweiten Beispiel, nicht systematisch kontrolliert verlaufen. Damit fehlt ein entscheidender Anker, um Selektivitäten zu kontrollieren und auf eine (möglichst bevölkerungsrepräsentative) Grundgesamtheit hochzurechnen. Daher ist die Belastbarkeit der erzielten Befunde nicht einzuschätzen. Eine sehr große realisierte Stichprobe wie im ersten Beispiel ist hier ein gewisser Vorteil. Mit wachsenden Fallzahlen und damit steigender Ausschöpfungsquote einer (unbekannten) Grundgesamtheit nehmen Selektivitätsverzerrungen tendenziell ab. Andererseits bildet die Stichprobe immer noch einen Bruchteil der Selbstständigen in Deutschland ab. Darüber hinaus ist wegen des eingesetzten Schneeball-Verfahrens damit zu rechnen, dass die Daten für verschiedene Untersuchungseinheiten miteinander korrelieren: Einladungen zur Teilnahme an der Befragung werden eher an Personen weitergeleitet, die der eigenen Person ähneln. Solche Cluster-Probleme lassen sich schwer kontrollieren, wenn der Verlauf der Empfehlungs-Ketten nicht nachvollzogen werden kann. In allen Beispielen wird nur ein verhältnismäßig kurzer Fragebogen eingesetzt. Hier tritt ein wesentlicher Nachteil des CAWI-Modus zutage – um die in digitalen Umgebungen relativ kurze Aufmerksamkeitsspanne von Befragten nicht zu überfordern, sollte die Beantwortung des Fragebogens nicht länger als 20 Minuten in Anspruch nehmen (vgl. ausführlicher Abschnitt 3.3).

2.7 Daten zu Absicherung und Lebenslagen im Alter

Seit 1986 wird im Auftrag des BMAS eine repräsentative Untersuchung zur Alterssicherung in Deutschland (ASID) durchgeführt (zuletzt im Jahr 2019), deren Ergebnisse in den Alterssicherungsbericht der Bundesregierung einfließen. Gegenstand der Untersuchung sind die Lebenssituation sowie Art, Höhe, Verteilung und Determinanten der Einkommen der Bevölkerung ab 60 Jahren auf der Ebene von Personen und Ehepaaren. Den Kern der Befragung bildet die Erfassung von 26 Einkommensarten, darunter verschiedene Rentenarten, aus denen sich das Alterseinkommen speisen kann. Die Datenerhebung basiert auf einer bei den Einwohnermeldeämtern gezogenen repräsentativen Adressenstichprobe.⁷ ASID ist als sequenzielle Methoden-mix-Studie (Mixed-Mode-Design) konzipiert. In der ersten Stufe wird eine schriftliche Befragung durchgeführt. Personen, die in der schriftlichen Phase nicht reagiert und verweigert haben, werden in der mündlichen Feldphase durch Interviewer kontaktiert und gegebenenfalls persönlich befragt. Die Wahl dieses Ansatzes begründet sich daraus, dass einige Personengruppen – darunter Selbstständige – schriftlichen Befragungen gegenüber erfahrungsgemäß aufgeschlossener sind als mündlichen, so dass durch ein Angebot verschiedener Befragungsmethoden eine höhere Ausschöpfung erzielt werden kann. Zudem konzentriert sich die Befragung auf sensible Einkommensfragen, über die gegenüber Dritten in Deutschland ungern Auskunft erteilt wird.

Im Jahr 2019 wurden im Rahmen von ASID etwas mehr als 30.000 Personen befragt. Die Ergebnisse lassen sich sehr gut hochrechnen, weil die Auswahlgesamtheit die Zielpopulation weitestgehend vollständig abbildet. Infolge der großen realisierten Stichprobe sind auch Selbstständige in nennenswerter Zahl enthalten. Heien und Krämer (2021) zeigen auf Grundlage von ASID 2019, dass 13 Prozent der Männer und fünf Prozent der Frauen ab 65 Jahren zuletzt selbstständig waren oder aktuell noch selbstständig sind. Dabei stellt sich die Gruppe hinsichtlich ihrer Einkommens- und Erwerbssituation als sehr heterogen dar. So beziehen etwa verkammerte Freiberufler deutlich mehr Erwerbseinkommen und Alterssicherungsleistungen als andere Gruppen von Selbstständigen wie Landwirte oder Handwerker.

Im Hinblick auf die Konzeption der geplanten Studie zur sozialen Lage von Selbstständigen ist das Beispiel ASID – trotz einer sehr hohen methodischen und Erhebungsqualität – allerdings nur von begrenztem Nutzen. Insbesondere eignet sich der Zugangsweg über die Einwohnermeldeämter nicht für eine spezifische Untersuchung zu selbstständigen Erwerbstätigen, da dort keine brauchbaren Daten vorliegen, um diese Personengruppe zu identifizieren. Ein Screening zur Identifikation der Zielgruppe ausgehend von entsprechendem bevölkerungsrepräsentativem Adressenmaterial im schriftlichen oder persönlichen Modus wäre angesichts der relativ kleinen Zielpopulation in hohem Maße unwirtschaftlich. Des Weiteren ist das ASID-Erhebungsprogramm auf ein sehr spezielles Erkenntnisinteresse ausgerichtet und primär auf Personen im Rentenalter zugeschnitten, während die durchzuführende Befragung auf Personen im Haupterwerbsalter fokussieren soll. Zur Erfassung der Qualität der Altersvorsorge von Erwerbstätigen sind die enthaltenen Items daher wenig geeignet.

Anders verhält sich dies bei Befragungen zur Altersvorsorge von Personen im Erwerbsalter. Diese können einen Bezugspunkt darstellen, wenn es darum geht, im Rahmen der geplanten Studie Befragungselemente zur Erfassung von Vermögensbildung und der sozialen oder privaten Vorsorge von Selbstständigen zur Einkommenssicherung im (Renten-)Alter zu entwickeln (vgl. Abschnitt 4.5). Eine solche Befragung liegt mit der SAVE-Studie („Sparen und Altersvorsorge in Deutschland“) vor. Hierbei handelt es sich um eine repräsentative Panelerhebung, die von 2001 bis 2013 durchgeführt wurde. Zuletzt wurde in SAVE das Vorsorgeverhalten von rund 1.430 Personen erfasst. Darunter

⁷ Dies ist ein „best practice“-Ansatz für Erhebungen, bei denen die Grundgesamtheit aus Personen in Privathaushalten besteht.

befanden sich, grob dem Anteil der Selbstständigen an allen Erwerbstätigen entsprechend, allerdings in jeder Welle weniger als 100 selbstständig Erwerbstätige (Börsch-Supan et al. 2009). Die Datenbasis war damit zu schmal, um robuste spezifische Aussagen zum Vorsorgeverhalten der Zielgruppe zu treffen. Nicht nur die Datenbasis, sondern auch der im Erhebungsprogramm reflektierte institutionelle und regulatorische Rahmen sind veraltet. Bei der Übernahme von Items aus SAVE wäre darauf zu achten, ob eventuell Adjustierungen bzw. Aktualisierungen vorzunehmen sind.

Eine andere Befragung zum Altersvorsorgeverhalten von Erwerbstätigen in Deutschland ist die Studie „Verbreitung der Altersvorsorge 2019“ (AV 2019). Im Mittelpunkt steht hier der aktuelle Erwerb von Anwartschaften in den drei Säulen der Alterssicherung: Gesetzliche Rentenversicherung (GRV), betriebliche Altersversorgung (BAV) sowie private Altersvorsorge, hierbei insbesondere Riester-Verträge. Allerdings beschränkt sich die Erhebung auf sozialversicherungspflichtig Beschäftigte (im Alter von 25 bis unter 65 Jahren). Im Fragenprogramm fehlen daher Items zu Vorsorgeformen, die bei der sich unter anderen Rahmenbedingungen vollziehenden Altersvorsorge von Selbstständigen eine besonders wichtige Funktion übernehmen. Dennoch können die in AV 2019 enthaltenen Fragen bei der Erarbeitung des Fragebogens für die geplante Studie ein Bezugspunkt sein. Schließlich sind zumindest bei hybrid Erwerbstätigen auch die ersten beiden Säulen der Alterssicherung mit in den Blick zu nehmen.

Der Zugang zu den Zielpersonen erfolgt in AV 2019 über eine disproportional geschichtete Stichprobe, die aus der – die vollständige Grundgesamtheit repräsentierenden – Beschäftigtenstatistik gezogen wird. Dieser Zugangsweg ist aber, wie oben (Abschnitt 2.5) im Kontext der IAB-Studien näher erörtert, bei Selbstständigen ausgeschlossen. Die Datenerhebung für die AV 2019 erfolgte mittels einer kombinierten schriftlichen Befragung, in der sowohl ein Papier- als auch ein Onlinefragebogen eingesetzt wurden. Die Entscheidung für diesen Erhebungsmodus war allerdings durch finanzielle Gründe bedingt. Im Ergebnis kam es zu systematischen Antwortausfällen bei Gruppen, die sich mit einem schriftlichen Format schwer tun. Tendenziell sind dadurch Personen, bei denen auch soziale Problemlagen zu vermuten sind, im Ergebnis unterrepräsentiert – und eine Hochgewichtung schwach besetzter Gruppen kann Zufallsfehler bei zugrundeliegenden geringen Fallzahlen im Kern nicht ausgleichen. Die Studie zeigt auch, dass bei einem schriftlichen Format der Einsatz eines „leichten“, wenig komplexen Fragebogens wesentlich ist. Bei einer Abfrage komplexer Sachverhalte, wie im Rahmen der geplanten Studie, die bereits zur Abgrenzung und Typisierung der selbstständigen Erwerbstätigkeit erforderlich sein wird (vgl. Abschnitt 4.2), stößt das Format schnell an Grenzen.

3. Möglichkeiten für die Anlage einer neuen Erhebung bei Selbstständigen

3.1 Integration in laufende Befragungen

Der Literaturüberblick zeigt, dass bereits eine Reihe von Befragungsdaten vorhanden sind, anhand derer sich erste Einschätzungen zu der sozialen Lage von Selbstständigen in Deutschland gewinnen lassen. Hierbei sind zwei Kategorien zu unterscheiden: (1) Erhebungen, die eine allgemeine Bevölkerungstichprobe umfassen, wie SOEP und PASS, (2) spezifische Erhebungen ausschließlich bei selbstständig Tätigen, wobei sich deren Grundgesamtheit zum Teil auf einzelne Teilgruppen, insbesondere Solo-Selbstständige, beschränkt.

Dagegen werden amtliche Daten, wie der Mikrozensus oder die administrativen Daten der Bundesagentur für Arbeit (BA), in diesem Zusammenhang eher wenig genutzt. Dies liegt im Fall des Mikrozensus vor allem am Fragenprogramm, das sich nur bedingt für die Operationalisierung des multidimensionalen Konstrukts der sozialen Lage eignet. So fehlt insbesondere eine präzise Erfassung der Einkommen, und Angaben zur Altersvorsorge und Krankenversicherung werden nur begrenzt erhoben. Im Fall der BA-Daten kommt zum Tragen, dass selbstständig Tätige in Deutschland überwiegend nicht in die Sozialversicherungen integriert sind. Darum werden, anders als bei abhängig Beschäftigten, erwerbsbiografische Informationen auf administrativer Ebene erheblich unvollständiger erfasst.

Die Nutzung von bevölkerungsrepräsentativen Erhebungsdaten hat den Vorteil, dass sich Befunde für selbstständig und nicht selbstständig Beschäftigte einander gegenüberstellen lassen. Spezifische Problemlagen, die mit der Selbstständigkeit verbunden sind, können so besser erkennbar werden. Wegen des relativ geringen Anteils von Selbstständigen in der Bevölkerung laufen empirische Auswertungen auf dieser Grundlage, wie in Abschnitt 2.1 im Zusammenhang mit dem SOEP näher ausgeführt, leicht in Fallzahlenprobleme; belastbare Aussagen für einzelne Teilgruppen innerhalb dieser in sich stark heterogenen Zielgruppe lassen sich auf einer schmalen Basis an Beobachtungen nur eingeschränkt gewinnen.

Eine Möglichkeit zur Überwindung dieser Beschränkung ist, eine laufende bevölkerungsrepräsentative Erhebung um eine zusätzliche, speziell gebildete Teilstichprobe von Zielpersonen – im vorliegenden Kontext also selbstständig Erwerbstätige – aufzustocken bzw. zu ergänzen. Für eine spezifische Studie zu einer bestimmten Thematik oder Bevölkerungsgruppe kommt dabei eine vollständige Übernahme des Erhebungsprogramms der Referenzstudie schon aus wirtschaftlichen Gründen kaum in Frage. So kostet im SOEP die Durchführung eines vollständigen regulären Interviews rund 300 Euro. Zudem sind zur Erreichung der Studienziele in der Regel nicht alle regulär abgefragten Variablen von Belang. Andererseits dürften im regulären Erhebungsinstrument Items fehlen, die für die Erfüllung der spezifischen Studienziele essenziell sind. Dies betrifft im vorliegenden Kontext etwa detaillierte Abfragen zur sozialen und privaten Absicherung gegen eine Reihe elementarer Lebensrisiken. Eine Verkürzung des regulären Fragebogens kann dann den notwendigen Spielraum schaffen, um diese Fragen aufzunehmen, da der zeitliche Rahmen des Interviews nicht beliebig ausgeweitet werden kann.

Wird das Fragenprogramm für eine an eine bevölkerungsrepräsentative Erhebung andockende Ergänzungsstichprobe ausgeweitet, sollten die für die Erfüllung der Studienziele wesentlichen zusätzlichen Fragen zumindest auch an die Zielpersonen gerichtet werden, die in der regulären Stichprobe ohnehin enthalten sind. Für diese Personen liegen bei einer Panel-Studie umfangreichere

Hintergrunddaten vor, und die angestrebte Anzahl von Beobachtungen für Zielpersonen kann so mit einer kleineren Ergänzungsstichprobe erreicht werden. Es kann jedoch auch nützlich sein, die entwickelten zusätzlichen Fragen auch an (ausgewählte) andere Personen in der regulären Stichprobe zu richten. Auf diese Weise lassen sich die Befunde für die Zielpopulation (im vorliegenden Kontext Selbstständige) mit den Befunden für eine relevante Vergleichsgruppe (im vorliegenden Kontext abhängig Beschäftigte) vergleichen und Besonderheiten erkennen, die die Zielpopulation charakterisieren.

Erweiterungen bzw. Anpassungen des Fragenprogramms in einer regelmäßig laufenden Erhebung erfordern allerdings erfahrungsgemäß einigen zeitlichen Vorlauf. Weil die Befragten nicht überfordert werden dürfen, stehen zusätzlich in die Haupterhebung einzuführende Items in Konkurrenz zu anderen Items. Dies begrenzt zum einen deren möglichen Umfang. Zum anderen müssen Wünsche für neue Items frühzeitig an die für die Erhebung Verantwortlichen herangetragen werden, damit diese im in der Regel langwierigen Auswahl- und Redaktionsprozess bei der Erstellung des Fragebogens berücksichtigt werden können. Die denkbare Alternative, in der Haupterhebung enthaltene Personen außerhalb dieser für eine spezifische Studie gesondert zusätzlich zu befragen, schafft zwar größere Spielräume für die Gestaltung des Fragebogens. Sie birgt aber das Risiko einer niedrigen Teilnahmebereitschaft, so dass nur eine selbstselektierte Stichprobe erreicht wird; bei Panel-Studien kann es sogar zu einem unerwünschten, die Qualität der Haupterhebung belastenden permanenten Ausscheiden von Untersuchungseinheiten kommen.

Insgesamt gesehen hat die Vorgehensweise, eine Ergänzungsstichprobe in eine bereits etablierte laufende Erhebung zu integrieren, einige Vorteile, bringt jedoch auch Verluste an Gestaltungsfreiheit mit sich. Dies betrifft Umfang und Ausrichtung des Fragebogens, den zeitlichen Rahmen einer Studie, der sich an den regulären Abläufen der laufenden Haupterhebung ausrichten muss, sowie die Abgrenzung der in die Ergänzungsstichprobe aufzunehmenden Untersuchungseinheiten. Damit sich Daten aus einer Ergänzungsstichprobe mit selbstständig Tätigen mit den Daten für selbstständig Tätige in der Hauptstichprobe konsistent zusammenführen lassen, muss „Selbstständigkeit“ in beiden Teilstichproben gleich operationalisiert sein. Die Auswahl der Untersuchungseinheiten für die Ergänzungsstichprobe muss demnach in der Tendenz anhand von Merkmalen gebildet werden, die für die Einheiten in der Hauptstichprobe ohnehin erfasst werden – es sei denn, es gibt Möglichkeiten, das laufende Erhebungsprogramm an alternative Kriterien zur Abgrenzung des Personenkreises in der Ergänzungsstichprobe anzupassen oder gegebenenfalls entsprechend zu erweitern.

Mehr Flexibilität bei der Gestaltung der Auswahlmerkmale für eine Ergänzungsstichprobe besteht, wenn sich mit im regulären Erhebungsprogramm enthaltenen Variablen verschiedene Definitionen umsetzen lassen. Beispielsweise werden bei Auswertungen des SOEP Selbstständige oft als Personen definiert, die nach eigenen Angaben in ihrer aktuellen beruflichen Haupttätigkeit selbstständige Tätigkeiten ausüben; mithelfende Angehörige werden nicht dazugezählt (vgl. etwa Bonin et al 2022). Jedoch erfasst das SOEP auch eine Anzahl von Angaben zur beruflichen Stellung für bis zu drei Nebentätigkeiten. Darunter sind die Kategorien „Selbstständige/Freiberufler“ und „Mithelfende Familienangehörige“. Somit könnte auch eine weiter gefasste Definition des Status „selbstständig“ unter Einschluss hybrider Arbeitsverhältnisse gewählt werden. Entsprechend ließe sich die Bildung einer Ergänzungsstichprobe im SOEP auf einer enger oder weiter gefassten Definition des Selbstständigen-Begriffs basieren.

Die im Vergleich zur Ergänzung einer laufenden Erhebung deutlich flexiblere Alternative ist die Durchführung einer eigenständigen Datenerhebung. Deren Gestaltung lässt sich voll auf die im Zentrum des Erkenntnisinteresses stehende Teilpopulation ausrichten – im vorliegenden Kontext also auf Erwerbstätige, die (auch) selbstständig arbeiten. Auch bei dieser Vorgehensweise ist es sinnvoll, sich beim Fragenprogramm an etablierten Erhebungen zu orientieren, soweit die zu

untersuchenden Sachverhalte darin abgebildet sind. Die Erhebungsprogramme sowohl des SOEP als auch des PASS enthalten zahlreiche Items, die gut geeignet sind, verschiedene Dimensionen zur Charakterisierung der sozialen Lage von Erwerbstätigen zu operationalisieren. Beide Datenbasen werden für diesen Zweck auch genutzt. Die Verwendung bereits gut erprobter Items minimiert Risiken und schafft Vergleichsmöglichkeiten, die zum einen der externen Validierung oder Plausibilisierung der Befragungsergebnisse dienen können, und zum anderen der Herausarbeitung von zielgruppenspezifischen Besonderheiten durch Gegenüberstellung von Befunden zum gleichen Sachverhalt für Nicht-Selbstständige. Insbesondere könnten und sollten Items aus dem SOEP oder anderen bevölkerungsrepräsentativen Erhebungen, die in der Forschung zu sozialen Lagen im Allgemeinen und zu prekärer (selbstständiger) Erwerbstätigkeit im Besonderen genutzt werden, in das für eine eigenständige Befragung zur sozialen Lage von Selbstständigen zu entwickelnde Erhebungsinstrument übernommen werden.

3.2 Grundgesamtheit und Zugangswege zur Zielgruppe

Unabhängig davon, ob die Ergänzung einer laufenden bevölkerungsrepräsentativen Erhebung oder eine eigenständig durchgeführte Erhebung vorgenommen werden soll: Es müssen Zugänge zur Zielgruppe gefunden werden, die zur Bildung von Stichproben geeignet sind, welche ein repräsentatives Abbild der Verhältnisse in der Grundgesamtheit von Selbstständigen liefern.

Eine fundamentale Schwierigkeit hierbei ist, dass kein zentrales Register existiert, in dem das Universum von Personen erfasst ist, die in irgendeiner Form selbstständig tätig sind. Zwar hinterlässt ein sehr hoher Anteil der Zielgruppe, soweit nicht nur ganz geringfügig – oder in der Schattenwirtschaft – selbstständig gearbeitet wird, individuelle Spuren in Steuerdaten, insbesondere im Rahmen der Einkommensteuererklärung bei der Erklärung von Einkünften aus selbstständiger Arbeit oder von gewerblichen Einkünften, oder durch Abgabe von Umsatzsteuererklärungen. Entsprechende administrative Individualdaten sind für Erhebungszwecke aus Gründen des Datenschutzes jedoch nicht zugänglich. Somit müssen andere Zugangswege zur Zielgruppe beschritten werden. Die folgenden Abschnitte beleuchten drei verschiedene Ansatzpunkte, die dafür in Frage kommen.

3.2.1 Unternehmensdatenbanken

Ein offener, wenn auch indirekter Zugangsweg zu selbstständig Tätigen sind kommerziell betriebene – datenschutzkonforme – Unternehmensdatenbanken. Diese bilden die Grundgesamtheit der wirtschaftsaktiven Unternehmen in Deutschland so gut wie vollständig ab, weshalb sie als Basis für repräsentative Analysen zur Unternehmenslandschaft sehr gut geeignet sind. Hierbei stehen zwei kommerzielle Unternehmensdatenbanken miteinander in Konkurrenz:

- die Unternehmensdatenbank der Bisnode Deutschland GmbH.⁸ In der Bisnode-Datenbank werden aktuell verfügbare öffentliche Daten über Unternehmen in Deutschland, Österreich und der Schweiz in Kombination mit von einer Fachredaktion recherchierten Informationen und Kennzahlen dargestellt. Die Datenbank enthält 5,6 Millionen Firmenprofile samt Informationen zu 7,5 Millionen Führungskräften inklusive Position und Funktion im

⁸ Für einführende Informationen vgl. <https://www.bisnode.de/globalassets/germany/pdf-dokumente/factsheets-und-produktflyer/bisnode-firmendatenbank-broschue-premium.pdf> (zuletzt aufgerufen am 12. August 2022).

Unternehmen. Die Bisnode-Datenbank umfasst Kontaktdaten, die es ermöglichen, die Führungskräfte direkt anzusprechen;

- die Unternehmensdatenbank der Unternehmensgruppe Creditreform „markus“.⁹ Diese enthält ausführliche, verlässliche und aktuelle Geschäftsinformationen zu handelsregisterlich eingetragenen deutschen, österreichischen und luxemburgischen Unternehmen sowie Entscheiderinnen und Entscheidern auf der ersten und zweiten Führungsebene. In Kombination mit dem „markus Gewereregister“, mit dem sich auch Unternehmen, die im Gewereregister eingetragen sind, recherchieren lassen, sind die aktuell wirtschaftsaktiven Unternehmen praktisch komplett recherchierbar. Zu den vorhandenen Informationen liegen Kontaktdaten und Ansprechpartnerinnen und Ansprechpartner vor, die eine Feldansprache ermöglichen.

Die Daten der Creditreform sind auch die Grundlage für das Mannheimer Unternehmenspanel (MUP).¹⁰ Das MUP ist ein seit 1992 am Leibniz-Zentrum für Europäische Wirtschaftsforschung (ZEW) für Forschungszwecke aufbereiteter und gepflegter Panel-Datensatz für Unternehmen mit Sitz in Deutschland. Seit 2000 ist der Gesamtbestand der Unternehmen in Deutschland halbjährlich erfasst. Es handelt sich um die von der Fallzahl umfangreichste Panel-Unternehmensdatenbank in Deutschland außerhalb der amtlichen Statistik. Das MUP eignet sich wegen der Abbildung der Population der wirtschaftsaktiven Unternehmen und der großen Anzahl der enthaltenen unternehmensindividuellen Informationen nicht nur zur direkten empirischen Untersuchung mikroökonomischer Fragestellungen. Vielmehr wird es auch intensiv genutzt als Ziehungspool für repräsentative Unternehmensbefragungen, wie etwa das Mannheimer Innovationspanel (MIP) oder das IAB/ZEW-Gründungspanel, sowie für spezielle Erhebungen bei abhängig vom inhaltlichen Fokus der Untersuchung nach bestimmten Merkmalen ausgewählten Gruppen von Unternehmen.

Bei der Ziehung von Bruttostichproben können aufgrund der großen Fallzahl in der Ausgangsgesamtheit gut Schichtungen nach bestimmten Merkmalen, wie etwa Sektor und Mitarbeiterzahl, vorgenommen werden. Belastbare Auswertungen bezüglich der Grundgesamtheit lassen sich im Fall einer Schichtung gewinnen, indem die unterschiedlichen Auswahlwahrscheinlichkeiten der Untersuchungseinheiten durch Hochrechnungsfaktoren korrigiert werden, mit denen die ursprünglichen Randverteilungen der Auswahlmerkmale wiederhergestellt werden. Ebenfalls lassen sich durch Abgleich der Verteilungen von für die Grundgesamtheit beobachteten Merkmalen mit den Verteilungen dieser Merkmale in der realisierten Nettostichprobe eventuelle Verzerrungen infolge selektiver Teilnahme an der Befragung beurteilen und statistisch korrigieren.

Als Stichprobenrahmen für eine repräsentative Erhebung bei Selbstständigen eignen sich die skizzierten Unternehmensdatenbanken allerdings nur mit zwei wesentlichen Einschränkungen. Erstens ist die Organisationseinheit Unternehmen nicht deckungsgleich mit der angepeilten Untersuchungseinheit der Selbstständigen. Einerseits kann es vorkommen, dass dieselbe Person als Unternehmerin oder Unternehmer mit mehreren Unternehmen verbunden ist. Um dadurch zustande kommende Mehrfachzählungen zu eliminieren, müsste zunächst ein Namensabgleich vorgenommen werden. Dies läuft dem vom Unternehmen ausgehenden Aufbau der Datenbanken zwar entgegen, lässt sich aber prinzipiell realisieren; so lassen sich die im MUP vorhandenen Informationen flexibel

⁹ Für einführende Informationen vgl. https://www.creditreform.de/fileadmin/user_upload/central_files/docs/produkte/marktanalyse-kundendaten/kundenbindung-akquise/markus/creditreform-markus-broschuere.pdf (zuletzt aufgerufen am 12. August 2022).

¹⁰ Für einführende Informationen vgl. <https://www.zew.de/forschung/mannheimer-unternehmenspanel> (zuletzt aufgerufen am 12. August 2022).

über Unternehmen hinweg miteinander verknüpfen. Im Ergebnis lässt sich so die im Verhältnis zum Unternehmenspool kleinere Grundgesamtheit der unternehmerisch tätigen Personen annähern. Die Befunde einer sich anschließenden Erhebung müssen auf diese Basis hochgerechnet werden. Die hochgerechneten Ergebnisse gelten dann für die Gruppe der Selbstständigen, die in einer Beziehung zu den registrierten Unternehmen in Deutschland stehen.

Andererseits ist damit zu rechnen, dass ein Teil der in den Datenbanken erfassten Führungskräfte, etwa als Angestellte mit der Funktion des Geschäftsführers, keiner selbstständigen Arbeit nachgeht. Hierbei können die Übergänge zwischen beiden Kategorien fließend sein. So können Geschäftsführerinnen und Geschäftsführer einer Gesellschaft, die selbst an dieser beteiligt sind, in der Perspektive des Sozialversicherungsrechts entweder „Unternehmer“ oder „Arbeitnehmer“ im klassischen Sinn sein. Entscheidend ist, welcher Status im konkreten Einzelfall überwiegt. Hierfür wiederum ist die Höhe der Beteiligung am Unternehmen, mit dem ein Anstellungsverhältnis besteht, ein zentrales Kriterium; unterschieden wird zunächst zwischen Geschäftsführerinnen und Geschäftsführern mit Minderheits- und beherrschender Beteiligung. Die in den Unternehmensdatenbanken vorliegenden Informationen sind in der Regel nicht hinreichend differenziert, um feine Unterscheidungen dieser Art verlässlich vornehmen zu können. Ob tatsächlich ein Status als Selbstständiger gegeben ist, müsste daher bei einer Kontaktaufnahme mit in den Ziehungspool aufgenommenen Führungskräften gescreent werden. Dabei erkennbar werdende Diskrepanzen zwischen zunächst vermutetem und tatsächlichem Selbstständigen-Status müssten bei der Hochrechnung gegebenenfalls durch Anpassung der Gewichtungsfaktoren korrigiert werden.

Eine zweite, gravierendere Einschränkung ergibt sich daraus, dass die Firmendatenbanken nur die Unternehmen mit Eintrag im Handels- oder Gewerberegister abbilden. Damit ist die im inhaltlichen Zusammenhang der geplanten Untersuchung zentrale Gruppe der Angehörigen der Freien Berufe nur lückenhaft erfasst, weil für sie keine Pflicht zur Gewerbeanmeldung besteht. Aus demselben Grund dürften auch Personen, die sich nur sporadisch selbstständig betätigen, häufig nicht erfasst sein. Da ein wesentliches Ziel der vom BMAS geplanten Erhebung ist, Einschätzungen zu prekären sozialen Lagen von Personen mit hybriden Arbeitsverhältnissen zu gewinnen, die nur im Nebenerwerb, geringfügig oder sporadisch selbstständig tätig sind, ist dies eine bedeutsame Einschränkung. Um ein repräsentatives Abbild des Universums der Personen mit selbstständigen Tätigkeiten zu gewinnen, reicht der Zugangsweg über Firmendatenbanken allein keineswegs aus.

3.2.2 Bevölkerungsrepräsentative Zufallsstichprobe

Da die Grundgesamtheit der vom BMAS geplanten Befragung zur sozialen Lage von Selbstständigen auf der Personenebene und nicht auf der Unternehmensebene angesiedelt ist, lässt sich eine Befragung der Zielgruppe auch ausgehend von einer bevölkerungsrepräsentativen Zufallsstichprobe als Stichprobenrahmen aufbauen. Da die Grundgesamtheit der Selbstständigen – anders als die Grundgesamtheit der Unternehmen – in keiner verfügbaren Datenbank erfasst ist, muss in einem ersten Erhebungsschritt ein Screening vorgenommen werden. Dies dient der Feststellung, ob eine zufällig aus der Bevölkerung ausgewählte Interviewperson zur Zielgruppe der eigentlichen Befragung gehört, also gemäß einer vorzugebenden Statusabgrenzung überhaupt als Selbstständiger zu klassifizieren ist. Ist diese Bedingung erfüllt und die betreffende Person zudem bereit, an einer späteren weiterführenden Befragung mitzuwirken, wird sie in den Ziehungspool für die Haupterhebung aufgenommen. Im zweiten Erhebungsschritt werden dann ausschließlich Personen mit selbstständigen Tätigkeiten adressiert.

Bei dieser Vorgehensweise empfiehlt es sich, die aus der Bevölkerungsgrundgesamtheit gebildete Bruttostichprobe zur Durchführung der Erhebung nach vorzugebenden Merkmalen zu schichten. Damit lässt sich gewährleisten, dass bestimmte Kategorien selbstständiger Personen oder

selbstständiger Tätigkeiten mit geringer Inzidenz in der Grundgesamtheit, an denen aber aus inhaltlichen Gründen ein besonderes Erkenntnisinteresse besteht, in der realisierten Stichprobe in hinreichender Zahl vertreten sind.

Allerdings kann der für das vorgeschaltete Screening zu betreibende Aufwand schnell zu einem limitierenden (Kosten-)Faktor werden. Dies veranschaulicht ein einfaches Rechenbeispiel: Angenommen, ein Prozent der Bevölkerung betriebe eine bestimmte Form selbstständiger Tätigkeit (z.B. Crowdwork), und ein Viertel der Interviews der per Screening identifizierten Personen mit dieser Tätigkeit könnten schließlich auch realisiert werden.¹¹ Dann müsste, um 250 Nettofälle für diese Form der selbstständigen Tätigkeit zu realisieren, die gescreente bevölkerungsrepräsentative Zufallsstichprobe bereits 100.000 Personen umfassen. Diese Limitierung lässt sich nur überwinden, wenn zusätzliche externe Informationen vorliegen, die eine gezielte Ansprache von Personen ermöglichen, die gesuchte, aber in der Gesamtbevölkerung nur selten anzutreffende Eigenschaften aufweisen. Damit entfernt man sich allerdings von einer reinen Zufallsstichprobe, und zusätzliche Selektivitätsprobleme können ins Spiel kommen (vgl. Abschnitt 3.2.3).

Weil auf der Screening-Stufe hohe Fallzahlen bewältigt werden müssen, damit am Ende für differenzierte Auswertungen hinreichende Nettofallzahlen zur Verfügung stehen, muss der erste Erhebungsschritt effizient organisiert werden. Ein möglicher Ansatzpunkt dafür ist, das Screening in eine laufende Omnibus-Umfrage zu integrieren. Omnibus-Befragungen werden von kommerziellen Anbietern laufend betrieben und können in Form von computergestützten telefonischen Interviews (CATI), computergestützten persönlichen Interviews (CAPI) oder als computergestütztes Web-Interviewing (CAWI) online durchgeführt werden.

CAPI- und CAWI-Modus haben jedoch den Nachteil, dass sie sich relativ schlecht skalieren lassen. Befragungen im CAPI-Modus sind im Vergleich am aufwändigsten zu organisieren, so dass Omnibus-Befragungen in der Regel nur in einem langsamen Rhythmus, etwa monatlich, durchgeführt werden. Auch sind die Erhebungskosten pro Standardfrage und Fall im Vergleich der drei Modi am höchsten. Omnibus-Befragungen im CAWI-Modus sind dagegen am günstigsten. Sie stoßen bei der Skalierung allerdings an Grenzen, weil die Teilnehmerinnen und Teilnehmer normalerweise aus Online-Access-Panels rekrutiert werden. Hierbei handelt es sich um einen begrenzten Bestand an Adressen von potenziellen Teilnehmerinnen und Teilnehmern an Online-Befragungen, die kommerzielle Betreiber über online- und teils auch offline-Kanäle geworben haben und die mit Anreizen zur Teilnahme motiviert werden. Durch dieses Verfahren sind Personen mit starker Online-Affinität und Personen, die stärker auf die gesetzten Anreize reagieren, zudem tendenziell überrepräsentiert.

Die Durchführungsfrequenz bei Omnibus-Befragungen im CATI-Modus ist sehr hoch – in der Regel finden Omnibus-Läufe mehrmals wöchentlich statt. Die Auswahl der Befragten erfolgt standardmäßig nach einem systematischen Zufallsverfahren, welches sicherstellt, dass alle Personen mit Telefonanschluss im Haushalt die gleiche statistische Chance haben, in die Stichprobe zu gelangen. Die Arbeitsgemeinschaft deutscher Marktforschungsinstitute (ADM) hat dabei ein Stichprobenverfahren, das sogenannte ADM-Design, geschaffen, mit dem nicht nur im Telefonbuch eingetragene Personen berücksichtigt werden, sondern mittels zufallsgenerierter Telefonnummern auch Personen, die dies nicht sind. Zudem ist ein sogenannter Dual-Frame-Ansatz Standard, also die Einbeziehung eines gewissen Anteils von Mobilfunknummern. Durch diese Vorgehensweise können repräsentative Ergebnisse für die deutschsprachige Bevölkerung im Erwerbsalter gewonnen werden. Zudem lassen sich wegen der hohen Durchführungsfrequenz und der umfassenden Grundgesamtheit

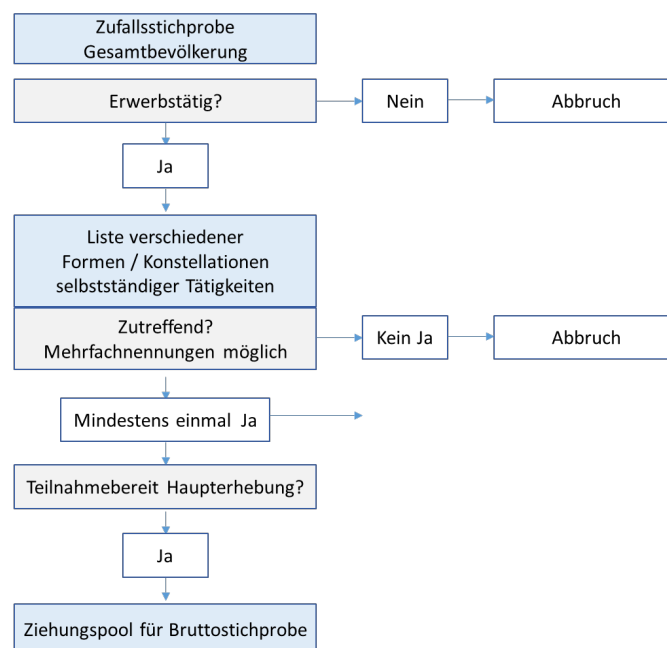
¹¹ Die Ausfallwahrscheinlichkeiten durch direkte Verweigerung einer Wiederkontaktingung und fehlende Mitwirkung bei Wiederkontaktingung kumulieren sich.

hohe Gesamtfallzahlen erreichen. Eine Omnibus-Befragung im CATI-Modus ist deshalb auch besonders geeignet, repräsentative Ergebnisse für kleinere Teilzielgruppen in der Bevölkerung zu erzielen. Dazu müssen die auf eine Teilzielgruppe ausgerichteten Fragen solange in die Läufe der Omnibus-Befragung eingespeist werden, bis eine ausreichende Gesamtfallzahl erreicht ist. Zum Beispiel müssen, um eine Nettostichprobe von 1.000 Fällen zu erreichen, bei einer Teilzielgruppe mit einem Prozent Anteil an der Bevölkerung 100.000 Personen gescreent werden. Bei 2.000 bis 2.500 realisierten CATI-Interviews pro Woche korrespondiert dies zu einer Feldzeit von 40 bis 50 Wochen. Dabei wird der für die Omnibus-Befragung verfügbare Bestand an verschiedenen Telefonnummern bei weitem nicht ausgeschöpft. Vielmehr ergibt sich die Grenze der Skalierbarkeit vor allem aus der für eine Untersuchung zur Verfügung stehenden Feldzeit.

Eine andere Begrenzung ist, dass im Rahmen einer CATI-Omnibusbefragung aufgrund des Mehrthemencharakters die Einschaltung von Fragen erfahrungsgemäß auf weniger als zehn begrenzt ist. Bei komplexeren Fragestellungen ist daher die gebräuchliche Vorgehensweise, im Rahmen der bevölkerungsrepräsentativen Mehrthemenumfrage identifizierte Zielpersonen – im vorliegenden Kontext also selbstständig Erwerbstätige – mit erklärter Bereitschaft für weitere Befragungen im Anschluss für eine umfangreichere spezielle Befragung erneut zu kontaktieren.

Dieses zweistufige Vorgehen hat auch den Vorteil, dass auf der Screening-Stufe eine Reihe elementarer sozio-demografischer Merkmale (üblich sind etwa Geschlecht, Alter, Bildungsabschluss, Haushaltgröße, Kinder im Haushalt, Familienstand, Haushaltseinkommen, Ortsgröße und Bundesland) erfasst werden, mit denen sich die in der zweiten Stufe realisierten Resultate auf die Gesamtbevölkerung hochrechnen lassen sowie mögliche Verzerrungen infolge selektiver Teilnahme an der Hauptbefragung beurteilt und gegebenenfalls statistisch korrigiert werden können. Dies ist im vorliegenden Kontext bedeutsam, weil die Verteilungen dieser Hintergrundvariablen für viele Typen von Selbstständigen, die für die Beurteilung der sozialen Lagen in dieser Erwerbstätigen Gruppe wesentlich sind (z.B. Selbstständige mit hybriden Arbeitsverhältnissen), nicht verlässlich bekannt sind.

Abbildung 3.1 Schema eines sparsamen Screening-Designs



Quelle: Darstellung des IZA.

Das Ablaufschema in Abbildung 3.1 illustriert, wie die Screening-Fragen im Rahmen einer CATI-Omnibusbefragung auf das absolute Mindestmaß begrenzt werden könnten, um den Ziehungspool für die Bruttostichprobe der zweiten Stufe möglichst sparsam anzulegen. Das dargestellte Design ist zudem flexibel in dem Sinn, dass Festlegungen bezüglich der Schichtung der Bruttostichprobe im Abschluss an die Screening-Phase getroffen werden können. Ergeben sich im Ziehungspool so niedrige Inzidenzen bei gewünschten Schichtungsmerkmalen, dass mit nicht belastbaren Fallzahlen in der schließlich realisierten Stichprobe zu rechnen ist, sollte auf ein Oversampling bezüglich dieser Merkmale verzichtet werden.

3.2.3 Selektive Zugangswege

Neben den beiden oben beschriebenen Zugangswegen, die auf eine – eventuell gezielt geschichtete – Zufallsauswahl der in die Erhebung einbezogenen Untersuchungseinheiten aus einer definierten Grundgesamtheit abzielen, gibt es weitere Möglichkeiten, selbstständig Tätige zu kontaktieren und für die Teilnahme an einer Erhebung zu rekrutieren. In Frage kommen vor allem:

- eine Kooperation mit Vereinen oder Verbänden, in denen sich Selbstständige organisieren. Diese Organisationen können allgemein über Befragungen informieren und für eine Mitwirkung werben, aber auch Mitglieder über den Adressverteiler individuell ansprechen und Zugangsdaten zu webbasierten Befragungen per E-Mail oder Papierfragebögen postalisch verteilen. Größere – organisationspolitische und datenschutzrechtliche – Hürden bestehen bei einer Weitergabe von Kontaktdaten an Dritte; eine direkte telefonische Kontaktierung durch Befragungsunternehmen etwa ist in der Praxis kaum möglich.
- die Werbung für eine Befragungsteilnahme über spezifische Social-Media-Kanäle oder andere, von der Zielgruppe besonders wahrgenommene Medien; eine Variante ist die Rekrutierung von Teilnehmerinnen und Teilnehmern über ein Schneeball-Verfahren, indem Zielpersonen aufgefordert werden, Zugangsinformationen an weitere Zielpersonen aus ihrem persönlichen Bekanntenkreis weiterzuleiten.
- die Rekrutierung über ein Online-Access-Panel, das heißt, einen kommerziell betriebenen Adressenpool aus potenziellen Teilnehmerinnen und Teilnehmern, wobei Befragte für ihre Teilnahme mittels spezieller Anreizsysteme belohnt werden. Soweit der Status „selbstständig“ in den hinterlegten Stammdaten erfasst ist, lässt sich dabei auch eine gezielte Stichprobe ziehen; die für die Mitglieder von Online-Access-Panels routinemäßige Bearbeitung von Fragebögen kann allerdings die Antwortqualität mindern. Eine Sonderform webbasierter Rekrutierung ist die Ausschreibung von Aufträgen zur Beantwortung von Fragebögen über sogenannte Mikrotask-Plattformen (wie „Clickworker“). Dieser Ansatz kommt naturgemäß insbesondere für Befragungen der speziellen Zielgruppe von Selbstständigen in Frage, die Crowdwork oder Gig Work nachgehen.

In der Literatur finden sich zwar einige Beispiele für Befragungen, die über Zugangswege dieser Art organisiert wurden. Jedoch ergeben sich dabei erhebliche Selektivitätsprobleme, bei denen sich Verzerrungen kumulieren: Erstens sind die ansprechbaren Personen keine Zufallsauswahl aus der relevanten Grundgesamtheit. So sind typabhängig nur bestimmte Selbstständige in einem für sie relevanten Verband organisiert oder nutzen Social Media. Zweitens entscheidet sich nur ein bestimmter Teil der angesprochenen Personen, an der Befragung teilzunehmen. Systematische Ausfälle durch fehlende Teilnahmebereitschaft ergeben sich zwar auch bei Studien, die auf zufällig aus der Grundgesamtheit gebildeten Bruttostichproben basieren. Jedoch sind die Rücklaufquoten bei den hier betrachteten Zugangswegen normalerweise markant niedriger, so dass sich das Problem der Selbstselektion in die Befragung noch verstärkt stellt.

Die doppelte Selektionsproblematik lässt sich mittels Gewichtungsfaktoren nicht ohne Weiteres befriedigend korrigieren. Die zugrundeliegende Grundgesamtheit ist nicht klar abgegrenzt, und die Treiber struktureller Differenzen zwischen Grundgesamtheit und realisierter Stichprobe lassen sich nicht hinreichend gut statistisch fassen. Erfahrungsgemäß genügt eine Hochrechnung anhand soziodemografischer Merkmale für eine Anpassung an die entsprechenden Randverteilungen gemäß Mikrozensus in der Regel nicht, um bevölkerungsrepräsentative Ergebnisse zu erhalten. Für die Bestimmung adäquater Gewichtungsfaktoren wären Informationen aus einer parallelen Befragung von Zielpersonen erforderlich, bei denen der erste Zugangsschritt nicht durch Selektivität geprägt ist. So wird teils zur Korrektur von Verzerrungen bei Online-Access-Panels verfahren, indem zusätzlich offline rekrutiert wird. Durch auf dieser Basis korrigierte Gewichtungsfaktoren kann das Kriterium der Repräsentativität zumindest für die Grundgesamtheit der Internet-Nutzerinnen und -Nutzer erfüllt werden.

Weil die geplante Untersuchung zur sozialen Lage von Selbstständigen bevölkerungsrepräsentative Befunde hervorbringen soll, kommen somit rein selektive Zugangswege zur Zielgruppe für die Umsetzung des Vorhabens nicht in Frage. Sie könnten allerdings neben einer repräsentativ angelegten Kernerhebung beschränkt werden. Dies kann, um eine höhere Zahl von Beobachtungen zu generieren, vor allem für wenig verbreitete, spezielle Konstellationen selbstständiger Tätigkeit sinnvoll sein. Einer besonderen Selektivität der realisierten Zusatzstichprobe wäre durch einen Vergleich mit der Kernerhebung analytisch nachzugehen. Die bei der Zusammenführung beider Datenbasen wahrscheinlich erforderliche Adjustierung von Gewichtungsfaktoren müsste anhand der Verteilungen zentraler Hintergrundfaktoren für die Personen in der Kernerhebung, die der Zielgruppe der Zusatzerhebung angehören, vorgenommen werden. Damit dies möglich ist, darf die Zahl der Zielpersonen in der Kernerhebung nicht zu klein sein. Für eine konsistente Zusammenführung von Kern- und selektiver Zusatzerhebung müssen beide zudem mit dem gleichen Erhebungsinstrument und im gleichen Erhebungsmodus erfolgen.

3.3 Erhebungsverfahren

Für standardisierte Befragungen in der empirischen Sozialforschung kommen mehrere Verfahren bzw. Erhebungsmodi in Frage. Jeder Modus hat jeweils unterschiedliche Vor- und Nachteile (Schnell 2019; Fowler 2014). Daher muss für jede Studie spezifisch über den am besten geeigneten Erhebungsmodus entschieden werden. Der Untersuchungsgegenstand (Art der Zielgruppe, Komplexität und Sensibilität des Erhebungsprogramms) und der Stichprobenrahmen, aber auch die benötigten wirtschaftlichen Ressourcen und eventuell vorgegebene zeitliche Fristen sind dabei zentrale Nebenbedingungen.

Folgt man dem in Abschnitt 3.1 erörterten Ansatz, eine spezifische Ergänzungsstichprobe von Personen aus der Zielgruppe an eine laufende bevölkerungsrepräsentative Erhebung anzudocken, ist damit der Erhebungsmodus praktisch vorgegeben. Er sollte mit dem Modus der Haupterhebung übereinstimmen, da mit einem Wechsel des Modus Verminderungen der Datenqualität einhergehen können. Diese können sich etwa aus unterschiedlichen Ausfallquoten („Unit-Nonresponse“), Unterschieden im partiellen Antwortausfall („Item-Nonresponse“) oder daraus ergeben, dass Personen bezüglich eines identischen Sachverhalts je nach Erhebungsmodus unterschiedlich antworten (Gensicke und Tschersich 2014; Jäckle et al. 2015; Lütke und Schupp 2017). Folgt man dem Modus der Haupterhebung, wären bei der Aufnahme einer zusätzlichen Selbstständigen-Stichprobe in das SOEP computergestützte persönliche Interviews (CAPI) und als Alternative selbst

ausgefüllte Papierfragebögen (Paper-and-Pencil Interviewing, PAPI) die Modi erster Wahl.¹² Dagegen wäre bei einer Stichprobenergänzung im Rahmen des PASS ein Mixed Mode aus CAPI und computergestützten telefonischen Interviews (CATI) zu wählen.

Wird eine eigenständige Erhebung konzipiert, ist dagegen Flexibilität bei der Moduswahl gegeben. Neben CAPI, CATI und PAPI kommen als weiterer gängiger Erhebungsmodus computerassistierte Web-Interviews (CAWI) in Frage. Dabei handelt es sich um Online-Umfragen, zu denen potenzielle Studienteilnehmerinnen und Studienteilnehmer üblicherweise per E-Mail eingeladen werden; prinzipiell sind jedoch auch andere Rekrutierungskanäle (Links in Social Media und auf Webseiten, postalisch, telefonisch) möglich. Teilnehmerinnen und Teilnehmer machen ihre Angaben dann selbstständig, indem sie einem über eine Webseite bereitgestellten Script folgen.

Insbesondere bei Befragungsinhalten, an denen die Untersuchungseinheiten kein starkes Eigeninteresse haben, oder bei Zielgruppen, die Befragungen grundsätzlich besonders skeptisch gegenüberstehen, wie es zumindest bei einigen Typen von Selbstständigen zu erwarten ist, kann die Beteiligung einer Interviewerin oder eines Interviewers die Mitwirkungsbereitschaft verbessern. Eine Möglichkeit hierzu ist die Durchführung einer mündlich-persönlichen Befragung, bei der die interviewende und die zu interviewende Person sich treffen und die interviewende Person die Antworten computergestützt (CAPI) oder in einem Papierfragebogen (PAPI) direkt aufnimmt.

Das mündlich-persönliche Befragungsverfahren hat drei zentrale Vorteile (Schröder 2015): Erstens können damit in der Tendenz Stichproben besser ausgeschöpft werden als mit anderen Modi. Zweitens lassen sich damit verhältnismäßig anspruchsvolle Inhalte erfassen. Vorab geschultes Interviewpersonal kann face-to-face schwierige Sachverhalte und Begriffe erläutern und eventuell auftretende Unklarheiten bei den Befragten ausräumen; implausible Antworten können durch Nachfragen im Gespräch korrigiert werden. Mittels visueller Hilfen wie Antwortlisten oder Kalenderdarstellungen sind auch komplexere Fragen und Filterführungen umsetzbar. Drittens können auch sehr umfangreiche Interviews durchgeführt werden. Im persönlichen Kontakt können Aufmerksamkeit und Motivation der Befragten länger bewahrt werden; die Wahrscheinlichkeit, Fragen nicht zu beantworten oder das Gespräch komplett abbrechen, nimmt ab. Deshalb sind bei Studien mit standardisierten persönlich-mündlichen Gesprächen Interviewzeiten von einer Stunde noch gut realisierbar – und in der Praxis häufig anzutreffen. Diesen Vorteilen steht als größter Nachteil gegenüber, dass der Erhebungsmodus in der Umsetzung mit hohen Durchschnittskosten pro Interview verbunden ist.¹³ Der Kostenfaktor steht in der Praxis bei beschränktem Budgetrahmen häufig der Realisierung ausreichend großer Nettostichproben entgegen – so dass auf kostengünstigere Erhebungsverfahren ausgewichen wird.

Der CATI-Modus stellt eine kostengünstige Alternative dar, bei der sich die mit der Mitwirkung eines Interviewers oder einer Interviewerin verbundenen Vorteile zum Teil bewahren lassen. Insbesondere kann geschultes Interviewpersonal im Gesprächsverlauf bei Verständnisproblemen der Befragten helfen und auf offensichtlich nicht plausible Antworten hinweisen. Im Vergleich zum persönlichen Interview ist der CATI-Modus wegen der stärker anonymen Gesprächssituation sogar tendenziell weniger von Interviewer-Effekten oder sozial erwünschtem Antwortverhalten betroffen. Dem stehen

¹² Im SOEP kommt ein größeres Spektrum an Erhebungsmodi zum Einsatz, wobei die einzelnen Modi je nach Teilstichprobe ein unterschiedliches Gewicht haben. Insgesamt dominiert aber die Nutzung von CAPI und PAPI.

¹³ Ein anders gelagerter, auf die Datenqualität wirkender Nachteil können Interviewer-Effekte sein. Davon spricht man, wenn das Antwortverhalten der Befragten systematisch von bestimmten Merkmalen (wie dem Geschlecht) oder der Gesprächsführung des Interviewers (etwa den bei Rückfragen mitgeteilten Informationen) abhängt. Auch tendieren Befragte im persönlichen Gespräch mit einem Gegenüber eher zu sozial erwünschten Antworten. Diese potenziellen Nachteile des Erhebungsverfahrens lassen sich aber mit organisatorischen Maßnahmen bzw. geeigneten Fragetechniken minimieren.

jedoch auch Nachteile gegenüber. So ist es schwieriger, einen hohen Ausschöpfungsgrad zu erreichen. Zum einen sind bestimmte Personenkreise, insbesondere jüngere Zielpersonen, telefonisch inzwischen nur noch schwer überhaupt erreichbar, selbst wenn mit einem Dual-Frame-Ansatz sowohl Festnetz- als auch Mobilfunknummern angesprochen und zur Stichprobenbildung per sogenanntem Randomized Last Digit-Verfahren (RLD-Verfahren) zufällige Telefonnummern generiert werden. Darüber hinaus hat die Bereitschaft zur Mitwirkung an telefonischen Umfragen auch bei den kontaktierbaren Personen in den letzten Jahren erheblich abgenommen. Dem ersten Problem lässt sich letztlich nur durch adjustierte Gewichtungsfaktoren begegnen, dem zweiten durch die Motivation der Zielpersonen stärkende Instrumente. Dazu gehören etwa gute Aufklärung im Hinblick auf das öffentliche Interesse an einer Studie, Vertrauen schaffende Maßnahmen bezüglich der Seriosität des Auftraggebers und des Datenschutzes, oder eine Incentivierung, das heißt, die Aussicht auf monetäre oder nicht-monetäre Vorteile im Fall einer Teilnahme. Solche Maßnahmen sind zugleich allerdings auch immer eine Quelle zusätzlicher Selektivitäten.

Außerdem muss das Erhebungsinstrument im CATI-Modus einfacher und knapper gestaltet werden. Weil optische Unterstützungsmöglichkeiten fehlen, verbrauchen komplexe Filterführungen mehr Zeit; ohne Kalendarium lassen sich retrospektive Verlaufsinformationen nur begrenzt erfassen. Informationen, die der befragten Person nicht sofort präsent sind, sind – anders als bei einer face-to-face Gesprächssituation – kaum spontan recherchierbar. In der Tendenz neigen Befragte im CATI-Modus eher zu unentschiedenen Antworten („weiß nicht“) als im CAPI-Modus. Auch ist die Aufmerksamkeitsspanne am Telefon kürzer. Dem kann durch eine abwechslungsreiche Gestaltung des Erhebungsinstruments etwas entgegengearbeitet werden. Selbst bei starkem Interesse der Befragten am Gegenstand der Untersuchung sollte die Interviewdauer im CATI-Modus in der Regel 45 Minuten nicht übersteigen. Oft lässt sich schon nach etwa einer halben Stunde eine merkliche Zunahme der Wahrscheinlichkeit des Gesprächsabbruchs von Seiten des Befragten beobachten.

Eine grundlegende Alternative zu den Erhebungsmodi mit Einsatz von Interviewpersonal sind sogenannte selbstadministrierte Befragungen, das heißt Studien, bei denen die Zielpersonen nach Einladung zur Teilnahme einen Fragebogen selbst ausfüllen (Menold 2015). Dies hat den Vorteil, dass Interviewer-Effekte komplett ausgeschlossen sind. Vor allem aber bedeutet die Organisation ohne Interviewpersonal eine erhebliche Reduktion des Kostenaufwands pro Fall. Somit können mit einem gegebenen Budget größere Stichproben erreicht werden als im CAPI- oder CATI-Modus. Dem steht allerdings entgegen, dass durch den weniger engen Kontakt zu den Zielpersonen mit geringeren Teilnahmequoten zu rechnen ist. Selbstadministrierte Befragungen erfordern also im Vergleich größere Bruttostichproben, um auf die gleiche Nettofallzahl zu kommen.

Angesichts der schwächeren Teilnahmebereitschaft sind in diesem Zusammenhang Maßnahmen zur Motivation der angesprochenen Zielpersonen noch einmal wichtiger. Um die subjektiven Kosten, den Fragebogen allein zu bearbeiten, gering zu halten und um Verständnisproblemen vorzubeugen, muss dieser unkompliziert zugänglich, mit klaren Instruktionen versehen und einfach zu beantworten sein. Der Fragebogen darf zudem nicht zu lang sein. Außerdem muss mit persönlichen oder sensiblen Fragen (etwa zu Vermögen) zurückhaltend umgegangen werden, weil sie leicht einen kompletten Ausstieg aus der Befragung oder zumindest Item-Nonresponse provozieren können. Wegen dieser Randbedingungen erscheinen selbstadministrierte Erhebungen erheblich schlechter dazu geeignet, komplexe Sachverhalte zu erfassen, als durch Interviewpersonal durchgeführte Befragungen.

Eine Variante der selbstadministrierten Erhebung ist die schriftlich-postalische Befragung. Dabei erhalten Zielpersonen einen Papierfragebogen per Brief. Die Umsetzung erfordert also eine Liste mit Postadressen, unter denen Zielpersonen der Studie anzutreffen sind, und die eine Zufallsauswahl aus der Grundgesamtheit der relevanten Postadressen darstellen sollte. Diese Bedingung ist in der Praxis oft schwer zu erfüllen. Zum Beispiel lässt sich die Zielgruppe der Selbstständigen anhand von

Melderegisterdaten nicht gezielt ansprechen. Schriftlich-postalische Befragungen leiden auch unter dem Nachteil, dass die Möglichkeiten der Fragebogengestaltung stark eingeschränkt sind. Die Papierform schließt komplexere Filterführungen praktisch aus, und die Bearbeitung des Fragebogens sollte nicht länger als zehn Minuten dauern.

In der Praxis wird der schriftlich-postalische Modus zunehmend durch CAWI abgelöst (Schnell 2019). Ein zentraler Vorteil der webbasierten selbstadministrierten Erhebung im Vergleich zur Papierform ist, dass sich mit der digitalen Gestaltung des Erhebungsinstruments (etwa durch responsive Steuerelemente oder Visualisierungen) deutlich komplexere Fragen und Filterführungen realisieren lassen. Interaktive Ausfüllhilfen und automatische Konsistenz-Checks können zudem zu einer höheren Antwortqualität beitragen. In diesen Punkten weist CAWI gegenüber PAPI ähnliche Vorteile auf wie CAPI gegenüber CATI. Ein weiterer, bei der Entscheidung für CAWI oft ausschlaggebender Vorteil ist, dass sich in diesem Modus große Fallzahlen sehr kosteneffizient und schnell realisieren lassen. Die benötigte Feldzeit ist normalerweise um einiges kürzer als bei den anderen Verfahren zur Durchführung standardisierter Interviews.

Das webbasierte Erhebungsverfahren hat jedoch auch Nachteile. So lassen sich im CAWI-Modus nach wie vor keine bevölkerungsrepräsentativen Ergebnisse erzielen. Die Internetreichweite ist mittlerweile zwar hoch, aber es fehlt ein verlässliches zentrales Register von Kontaktdaten. Teile der Bevölkerung sind trotz vorhandenem Internetzugang nur schlecht für Befragungszwecke erreichbar; spezifische Grundgesamtheiten nach sozio-demografischen Merkmalen sind deswegen nur eingeschränkt abgrenzbar. Dennoch lässt sich für einige Bevölkerungsgruppen, wie Jugendliche oder Personen mit hohem Bildungsgrad, inzwischen im CAWI-Modus ein hohes Maß an Repräsentativität innerhalb dieser Gruppen erreichen. Im Hinblick auf die heterogene Gruppe der Selbstständigen dürfte dies derzeit allerdings noch nicht gewährleistet sein. Darüber hinaus leiden webbasierte Interviews wegen der starken Konkurrenz um Aufmerksamkeit im Internet erfahrungsgemäß unter einem niedrigen Ausschöpfungsgrad, so dass zusätzliche Selektivitätsprobleme entstehen können. Um Personen zur Teilnahme zu motivieren, werden häufiger substantielle Incentivierungen eingesetzt. Dies kann die Datenqualität verschlechtern, etwa weil der digitale Fragebogen nur oberflächlich bearbeitet wird, um eine in Aussicht bestellte Belohnung zu erhalten.¹⁴ Außerdem neigen Befragte in der Anonymität des Internets zur Artikulation extremerer Meinungen. Während der Einsatz von Interviewpersonal das Risiko eines sozial erwünschten Antwortverhaltens in sich trägt, kann somit das Pendel beim CAWI-Modus in die andere Richtung zu weit ausschlagen.

Schließlich beschränkt die kurze Aufmerksamkeitsspanne im Internet die Länge des Fragebogens. Bearbeitungsdauern im CAWI-Modus von mehr als 20 Minuten werden als kritisch angesehen (Gensicke und Tschersich 2014). Somit sind zwar längere Erhebungen durchführbar als im schriftlich-postalischen Modus, aber nur deutlich kürzere Erhebungen als im CATI- oder gar CAPI-Modus. Zwar besteht prinzipiell die Möglichkeit, dass zufällig ausgewählten Personen nur Teile des Fragebogens gezeigt werden und zufällig ausgewählten anderen Personen andere Teile. Für diesen Split muss aber eine größere Stichprobe realisiert werden, und die kohärente Zusammenführung gesplitteter Daten ist nicht trivial.

In der Praxis werden in einigen Fällen verschiedene Befragungsmodi auch gemischt. So ist es im Kontext von selbstadministrierten Erhebungen möglich, den Zielpersonen sowohl den PAPI- als auch den CAWI-Modus anzubieten. Dazu kann etwa in einem postalischen Anschreiben zur Übermittlung

¹⁴ Durch Aufzeichnung des Nutzerverhaltens während der Beantwortung der Fragen lassen sich Hinweise auf ein solches Verhalten gewinnen. Der notwendige Hinweis auf eine solche Überwachung kann jedoch auch eigentlich motivierte potenzielle Teilnehmerinnen und Teilnehmer abschrecken.

eines auf Papier auszufüllenden Fragebogens ein Link oder ein QR-Code zum Aufrufen einer online-Version dieses Fragebogens übermittelt werden. Andererseits kann den Adressaten einer online-Erhebung angeboten werden, dass der Fragebogen auf Anforderung in Papierform postalisch zugestellt wird. Auf diese Weise erhalten die Adressaten zwei Antwortoptionen zur Auswahl, womit unterschiedlichen Präferenzen besser Rechnung getragen werden kann. Dies kann die Teilnahmebereitschaft verbessern und damit die Ausschöpfung der Bruttostichprobe verbessern. Allerdings bleiben bei diesem Vorgehen die Beschränkungen der beiden miteinander kombinierten Befragungsmodi erhalten. So lässt sich die postalische Erstansprache einer repräsentativen Bruttostichprobe im vorliegenden Kontext nicht umsetzen, weil eine dafür geeignete Datenbank mit vollständigem Adressmaterial für die Zielgruppe der Erwerbstätigen, die ganz oder teilweise selbstständige Tätigkeiten ausüben, fehlt. Wenn im PAPI- und CAWI-Modus der gleiche Fragebogen eingesetzt wird, lassen sich die eigentlichen Vorteile eines digitalen Erhebungsinstruments, wie responsive Elemente oder Visualisierungen, nicht ausschöpfen.

Eine andere Form der Kombination von Befragungsmodi ist, die Befragung zu splitten, indem ein Teil der zu erfassenden Daten mit einem mündlich-persönlichen Verfahren erhoben wird und ein anderer Teil der Daten per selbstadministrierter Erhebung. Dazu werden im ersten Schritt Informationen im CAPI- oder CATI-Modus erhoben. Anschließend werden im sogenannten „drop-off-Verfahren“ Zugangsdaten zu einem webbasierten Fragebogen übermittelt oder ein Papierfragebogen hinterlassen, um zusätzliche Daten bei Zielpersonen zu erheben, die noch weiter auskunftsbereit sind. Dieses Vorgehen hat den Vorteil, dass in den verschiedenen Teilen der Befragungen an den gewählten Modus angepasste Erhebungsinstrumente verwendet und deren jeweilige Stärken ausgeschöpft werden können. So lassen sich etwa sensible Abfragen, die im mündlich-persönlichen Verfahren durch Interviewer-Effekte gestört werden könnten, in einen selbst auszufüllenden Fragebogen verlagern; durch teilweise Nutzung des CAWI-Modus können komplexe Sachverhalte erhoben werden, für deren korrekte Abfrage eine interaktive oder optische Unterstützung hilfreich ist. Durch verschiedene Befragungsmodi kann auch die Aufmerksamkeitsspanne der Zielpersonen verlängert werden, so dass in der Summe verschiedener Befragungsteile ein umfangreicheres Erhebungsprogramm realisiert werden kann.

Diesen Vorteilen steht jedoch als gravierender Nachteil entgegen, dass bei dem im drop-off-Verfahren zu realisierenden Befragungsteil mit erheblichen Ausfällen an Teilnehmenden zu rechnen ist, selbst wenn diese im persönlich-mündlichen Modus einer weitergehenden Befragung zustimmen. Dem muss durch eine größere Ausgangsstichprobe Rechnung getragen werden, was den gesamten Erhebungsaufwand erhöht. Darüber hinaus sind Verzerrungen zu erwarten, weil es sich bei den Teilnehmenden der zweiten Stufe um eine selbst selektierte Gruppe handelt. So könnten im vorliegenden Kontext selbstständig Tätige mit prekärer soziale Lage eher bereit sein, weitere Auskünfte zu geben, um so auf diese aufmerksam zu machen. Es ist nicht gesichert, dass sich solche Verzerrungen durch Anwendung von Hochrechnungsfaktoren adäquat korrigieren lassen. Aufgrund dieser Nachteile sollten drop-off-Verfahren hauptsächlich dann eingesetzt werden, wenn es um Sachverhalte geht, die im mündlich-persönlichen Verfahren nicht ausreichend zu erfassen sind, wie insbesondere persönlich hochsensible Fragen. Daneben kommt es auch in Frage, um qualitative Einschätzungen zu Aspekten zu gewinnen, die sich am Rand der Kernfragestellung bewegen. Im vorliegenden Zusammenhang wurde ein Splitting der Befragung mittels drop-off-Verfahren jedoch von keinem der diesbezüglich interviewten Umfrageunternehmen empfohlen.

Der Vergleich der verschiedenen gängigen Befragungsverfahren zeigt, dass Länge und Komplexität des geplanten Fragenprogramms für die Wahl des geeigneten Erhebungsmodus zentral sind. Das folgende Kapitel umreißt die – umfangreichen – Befragungsinhalte, die für die vom BMAS geplante Untersuchung zur sozialen Lage von Selbstständigen abgedeckt werden müssen.

4. Befragungsinhalte

Die Inhalte der geplanten Datenerhebung sind an dem Ziel auszurichten, auf Grundlage einer geeigneten Datenbasis belastbare und differenzierte Einschätzungen zur sozialen Lage von Selbstständigen zu gewinnen. Insbesondere gilt es, durch geeignete Erhebungsinstrumente eine aussagekräftige und detaillierte Informationsbasis bereitzustellen, um die Frage zu beantworten, wie häufig prekäre Lebensverhältnisse bei Selbstständigen auftreten. Darüber hinaus soll es mit weiteren zu erhebenden Merkmalen möglich sein, die Gründe für prekäre Lebensverhältnisse bei Selbstständigen und deren Folgen – insbesondere für ihre soziale Absicherung und ihre Vorsorge für den Fall von Einkommenseinbußen, etwa bei Krankheit, im Alter oder bei wirtschaftlichen Krisen – näher zu beleuchten.

Die Inhalte der vom BMAS geplanten Erhebung müssen sich an den schon einleitend behandelten Antworten auf die Frage orientieren, wie sich prekäre wirtschaftliche und soziale Lebensverhältnisse für die Gruppe der Selbstständigen messen lassen. Demnach gilt es, einem multidimensionalen Konzept von Prekarität folgend, eine Reihe unterschiedlicher, potenziell kumulativ wirkender Faktoren zu operationalisieren und deren Prekarität begünstigende Ausprägungen anhand geeigneter Indikatoren zu erfassen (vgl. etwa Schulze Buschoff et al. (2017); Promberger et al. (2018)). Zu unterscheiden sind hierbei die

- Dimensionen der Prekarität der individuellen Beschäftigungslage (vgl. Tabelle 1.1), wobei in der Diagnose gegebenenfalls über mehrere (parallel ausgeübte oder zeitlich eng aufeinander folgende) Beschäftigungsverhältnisse hinweg aggregiert werden muss, und
- die Dimensionen prekärer Haushaltslagen, mit denen selbstständig Tätige konfrontiert sein können, selbst wenn ihre eigene Position in der Arbeitswelt nicht als prekär zu klassifizieren ist (vgl. Tabelle 1.2).

Das nachfolgend skizzierte Fragenprogramm umfasst entsprechend die Erhebung von Merkmalen sowohl auf der Individualebene als auch auf der Haushaltsebene. Diese decken für die befragte Person und den Haushalt eine Reihe von sozio-demografischen Basismerkmalen ab, mit denen sich Auswertungen getrennt für relevante Teilpopulationen (individuell etwa nach Alter oder Geschlecht; bezüglich der Haushalte etwa für verschiedene Familientypen) vornehmen lassen, aber auch zur Operationalisierung prekärer sozialer Lagen dienen können (etwa bezüglich der Wohnverhältnisse).

Auf individueller Ebene ist gemäß dem Ziel der vom BMAS geplanten Studie zunächst die aktuelle Erwerbssituation von Personen, die selbstständige Tätigkeiten ausüben, möglichst differenziert zu erfassen. Dabei sind insbesondere auch multiple selbstständige und hybride Tätigkeiten zu erheben. Ergänzend dazu sollten Angaben zum selbstständig betriebenen Geschäft oder Unternehmen gesammelt werden. Des Weiteren sind Informationen zur bisherigen Erwerbsbiografie der selbstständig Erwerbstätigen essenziell, insbesondere die Dauer und Stabilität früherer selbstständiger oder abhängiger Beschäftigungen seit dem Eintritt in das Erwerbsleben. Damit lassen sich Dynamiken erkennen, die aktuell prekäre Lagen von Selbstständigen begünstigen. Das vorgeschlagene Erhebungskonzept zur Natur der selbstständigen Tätigkeit ist dabei so konzipiert, dass eine aktuelle Selbstständigkeit als Anker fungiert; jedoch sollten auch Perioden selbstständiger Erwerbstätigkeit in der jüngeren Vergangenheit berücksichtigt werden. Damit lässt sich eine Untererfassung sporadisch ausgeübter Selbstständigkeit vermeiden und der Volatilität hybrider Erwerbskonstellationen Rechnung tragen.

In weiteren Themenkomplexen zum Einkommen, Vermögen und Vorsorgeverhalten ist vorgesehen, in das Fragenprogramm jeweils den Haushaltskontext einzubeziehen, um zu berücksichtigen, dass eine prekäre individuelle Erwerbslage durch den familiären Kontext aufgefangen werden kann, oder dass selbst nicht prekär Beschäftigte im Haushaltskontext mit einer prekären Lage konfrontiert sein können.

Schließlich wird angeregt, in das Fragenprogramm auch subjektive Einschätzungen, persönliche Einstellungen und Haltungen sowie Persönlichkeitsmerkmale aufzunehmen. Diese Größen können wesentliche Faktoren sein, die in einem systematischen Zusammenhang mit der Inzidenz und Schwere prekärer sozialer Lagen stehen. Eine solche Erweiterung des Fragenprogramms schafft zudem einen Ausgleich zu den anderen primär faktenorientierten Themenkomplexen – und kann wesentlich dazu beitragen, die Motivation der Befragten hoch zu halten und Abbruchrisiken zu senken.

Im Folgenden wird für die genannten Themenbereiche skizziert, welche inhaltlichen Dimensionen jeweils darunter erfasst werden sollten. Hingegen wird auf die Ausführung eines konkreten Fragenprogramms verzichtet. Diese wäre ebenso wie die Entwicklung einer im Hinblick auf die Minimierung von Abbrüchen und partiellen Antwortausfällen geeigneten Fragebogendramaturgie vom mit der vom BMAS zu vergebenden Studie beauftragten Anbieter in der ersten Phase der Projektdurchführung zu leisten. Dabei sollte bei der Entwicklung des Fragebogens, wo möglich, auf in anderen laufenden Befragungen bereits genutzte Items zurückgegriffen werden. Auf diese Weise eröffnen sich Möglichkeiten, die erzielten Befunde für die Zielpopulation von selbstständig Tätigen mit Befunden für andere Gruppen von Erwerbstätigen zu vergleichen.

4.1 Basismerkmale auf Individual- und Haushaltsebene

Die Erhebung individueller sozio-demografischer Basismerkmale kann sich weitgehend an etablierten Erhebungsinstrumenten orientieren. Neben Merkmalen wie Geschlecht, Alter und Familienstand erscheint es nicht erforderlich, einen detaillierten individuellen Migrationshintergrund mit der dazu notwendigen, sehr komplexen Fragenstruktur zu erheben. Vielmehr sollte es ausreichend sein, eine etwaige Zuwanderung der befragten Person sowie gegebenenfalls deren Zeitpunkt zu erfragen. Bei der Erfassung der Bildung und Ausbildung der befragten Person bietet es sich an, in laufenden Befragungen erprobte und etablierte Erhebungsinstrumente zu verwenden. Darüber hinaus erscheint es mit Blick auf die Zielgruppe der Selbstständigen sinnvoll, nach einer etwaigen selbstständigen oder unternehmerischen Tätigkeit in der Elterngeneration zu fragen.

Auf Haushaltsebene sollte erfasst werden, ob eine Partnerin oder ein Partner vorhanden sind, und ob weitere Personen, insbesondere Kinder, in der Bedarfsgemeinschaft leben. Als Dimensionen der sozialen Lage auf Haushaltsebene sollten einerseits Angaben zum verfügbaren Einkommen und andererseits zur Wohnsituation erfasst werden.

Die folgende Übersicht fasst zentrale Variablen zusammen, die als Basismerkmale auf der Individualebene und anhand der Angaben des Befragten für den Haushalt erfasst werden sollten.

Basismerkmale auf Individualebene

- Geschlecht, Alter, Familienstand
- Bildung und Ausbildung (höchster Schulabschluss, beruflicher Bildungsabschluss)
- Selbst nach Deutschland zugewandert? Falls ja, wann?
- Vater oder Mutter selbstständig/Unternehmer*in?

Basismerkmale auf Haushaltsebene

- Partner*in ja/nein, im Haushalt lebend
- Kinder ja/nein, Anzahl, im Haushalt lebend, Alter des jüngsten Kindes
- verfügbares Einkommen im Monat (nach Steuern und Sozialbeiträgen, aber vor anderen regelmäßigen Ausgaben)
- Laufende monatliche Verpflichtungen durch Warmmiete/Hausgeld, Versicherungen
- Wohnsituation: Wohnfläche, Eigentum

4.2 Erwerbssituation und Charakter der selbstständigen Erwerbstätigkeit

Eine an den Mikrozensus angelehnte Abfrage zur Erfassung der aktuellen Erwerbssituation der Befragten erscheint im hier betrachteten Rahmen zu komplex. Sie ließe sich zudem nicht konsistent mit einem zweistufigen Vorgehen mit sparsamer Filterung zum Screening einer bevölkerungsrepräsentativen Bruttostichprobe (vgl. Abbildung 3.1) verbinden.

Als Alternative wird vorgeschlagen, im Rahmen der geplanten Erhebung von einer aktuell ausgeführten selbstständigen Erwerbstätigkeit als Anker auszugehen und daran anknüpfend vertiefende Fragen zu deren Charakter zu stellen. Diese Fragen sollten sich insbesondere auf Umfang, Dauer und Art¹⁵ der selbstständigen Tätigkeit beziehen. Mehrere parallel – selbstständig oder abhängig beschäftigt – ausgeübte Tätigkeiten müssen gegebenenfalls in geeigneter Weise erfasst werden, damit bei hybriden Erwerbsmodellen zwischen Haupt- und Nebenerwerbstätigkeiten unterschieden werden kann.

Um schwankenden Arbeitszeiten und möglicherweise volatilen Erwerbskonstellationen Rechnung zu tragen, sollte daneben auch eine rückblickende zeitraumbezogene Erfassung selbstständiger Erwerbstätigkeit vorgenommen werden. Das hierfür verwendete Zeitfenster sollte nicht zu kurz gewählt werden, um die Inzidenz nur sporadisch und kurzzeitig ausgeübter selbstständiger Erwerbstätigkeit nicht zu unterschätzen. Andererseits sollte es auch nicht zu lang sein, um dem Problem vorzubeugen, dass Befragte in der Rückschau kurze, geringfügige Episoden mit selbstständiger Tätigkeit nicht oder nicht korrekt erinnern. Ein diese Kriterien erfüllender

¹⁵ Zur Erfassung der Art der selbstständigen Tätigkeit(en) sollte eine Liste mit wesentlichen Typen genutzt werden. Eine Typisierung, deren einzelne Elemente sich teils nicht wechselseitig ausschließen, könnte umfassen: Selbstständig und Arbeitgeber, Selbstständig ohne abhängig Beschäftigte (Solo-Selbstständig), Freiberufliche Tätigkeit, Selbstständig mit angemeldetem Gewerbe, Geschäftsführende Tätigkeit, Gig Work oder Crowdwork, Gelegenheits- oder Saisonarbeit.

Referenzrahmen wäre die Bezugnahme auf die Erwerbsverhältnisse während der letzten zwölf Monate vor dem Zeitpunkt der Befragung. Bei einer so basierten Messung würden auch eventuelle systematische Schwankungen in der Inzidenz selbstständiger Tätigkeiten im Jahresverlauf bzw. Saisonalitäten ausgeglichen. In dieser Hinsicht hat der vorgeschlagene Bezugsrahmen von zwölf Monaten Vorteile gegenüber denkbaren Alternativen wie Zeitfenstern von sechs oder 18 Monaten.

Erwerbssituation von Personen mit selbstständiger Tätigkeit

- Aktuell erwerbstätig, dabei ausschließlich oder zumindest teilweise selbstständig tätig?
- Wenn aktuell selbstständig tätig, welche Art der selbstständigen Tätigkeit (Mehrfachnennungen möglich), jeweils seit wann?
- Wenn aktuell nicht selbstständig tätig, innerhalb der letzten zwölf Monate selbstständig gewesen? Welche Art, jeweils seit wann, und seit wann nicht mehr? Gründe für Beendigung, parallel dazu auch abhängig beschäftigt gewesen?
- Wenn aktuell abhängig beschäftigt, seit wann? Wenn nicht aktuell, sondern in den letzten zwölf Monaten, seit wann, und seit wann nicht mehr?
- Umfang der selbstständigen Erwerbstätigkeit
- Durchschnittliche tatsächliche Arbeitszeit pro Woche aktuell sowie in den letzten zwölf Monaten
- Wenn aktuell oder in letzten zwölf Monaten hybrid erwerbstätig: jeweils Anteil der auf selbstständige Tätigkeiten insgesamt entfallenen Arbeitszeit
- Wenn zwei oder mehr Arten selbstständiger Tätigkeit, welche ist die mit der längsten, welche mit der zweitlängsten Arbeitszeit in den letzten zwölf Monaten? (Klassifikation in Haupt- und 1. Nebenselbstständigkeit), jeweiliger Anteil an der Arbeitszeit insgesamt

Angaben zum selbstständigen Geschäft

(möglichst differenziert nach Haupt- und 1. Nebenselbstständigkeit)

- Wenn Gewerbe angemeldet, seit wann?
- Wirtschaftszweig, Beruf, bei hybriden Tätigkeiten auch der abhängigen Haupttätigkeit
- Richtung der Entwicklung von (erwarteten) Umsätzen letzte drei, zwölf Monate, nächste zwölf Monate
- Richtung der Entwicklung der (erwarteten) Beschäftigung von Mitarbeiter*innen letzte drei, zwölf Monate, nächste zwölf Monate
- Aktuelle Wettbewerbsposition: Anzahl der Kunden, Auftraggeber, Bewertung des Wettbewerbsdrucks, zentraler Wettbewerbsparameter (Preis, Qualität, etc.)

Um Hinweise zu erhalten, in welchen Segmenten der Wirtschaft prekäre Selbstständigkeit besonders auftritt, welche Marktbedingungen und Geschäftslagen systematisch mit Prekarität verbunden sind, oder ob Gründer*innen besonders oft mit prekären wirtschaftlichen und sozialen Lagen konfrontiert

sind, sind über die Angaben zur Erwerbssituation auf der Personenebene hinaus auch Informationen zum im Rahmen der Selbstständigkeit betriebenen Geschäft von Belang. Da Selbstständige zum Teil mehrere Geschäfte parallel betreiben, können hierbei Abgrenzungs- oder Aggregationsprobleme auftreten. Deswegen sollte ein konkreter Bezug hergestellt werden. So könnten die erhobenen Angaben auf die selbstständige Tätigkeit mit dem höchsten Anteil an der aktuellen Arbeitszeit („Hauptselbstständigkeit“) bezogen werden, wenn möglich zusätzlich auch für die selbstständige Tätigkeit mit dem zweithöchsten Anteil („1. Nebenselbstständigkeit“).¹⁶

Die obenstehende Übersicht enthält die zentralen Variablen, die zur Erfassung der Inzidenz und des Charakters selbstständiger Erwerbstätigkeit bei den Zielpersonen erhoben werden sollten.

4.3 Einkommen

Zur Beurteilung der persönlichen Einkommensverhältnisse der selbstständig arbeitenden Zielperson ist es erforderlich, nicht nur deren gesamte Einkünfte aus selbstständiger Tätigkeit zu erfassen, sondern darüber hinaus auch andere regelmäßige persönliche Einkommen, vor allem solche aus einer eventuell ebenfalls ausgeübten abhängigen Beschäftigung. Dabei sollten nicht nur die Einkünfte im letzten Monat, sondern zusätzlich diejenigen im letzten Kalenderjahr abgefragt werden, um volatilen Erwerbskonstellationen und mithin schwankenden Einkommen Rechnung zu tragen. Darüber hinaus ist eine explizite Nachfrage nach der Volatilität der monatlichen Einkünfte sowohl aus selbstständiger als auch aus sonstigen persönlichen Quellen geboten – hohe (Einkommens-)Unsicherheit ist eine wesentliche Dimension, die zu prekärer Beschäftigung beiträgt.

Von zentraler Bedeutung ist es, die Einkommensverhältnisse im gesamten Haushalt zu erfassen; geringe persönliche Einkünfte bedingen nicht notwendigerweise eine prekäre individuelle soziale Lage, sofern diese mittels anderer Einkünfte durch die Bedarfsgemeinschaft aufgefangen werden können. Fragenstellungen nach Einkünften sollten daher weitgehend analog zu denjenigen nach persönlichen Einkünften mit Bezug auf eine Lebenspartnerin oder einen Lebenspartner in der Bedarfsgemeinschaft erfolgen, wobei gewisse Vergrößerungen aus Gründen der Datensparsamkeit sinnvoll sind. Zusätzlich sollte auch der Bezug der wichtigsten haushaltsbezogenen Sozialleistungen durch den Haushalt erfasst werden.¹⁷

Des Weiteren existieren etablierte Instrumente zur Erfassung des Sparverhaltens von Haushalten (etwa im SOEP, in der SAVE-Studie oder der Einkommens- und Verbrauchsstichprobe EVS), die in das Fragenprogramm eingebaut werden können. Diese sollten gute Anhaltspunkte dafür geben, wie angespannt die finanzielle Situation des Haushalts ist und ob – gegebenenfalls in welcher Höhe – monatliche Einkünfte regelmäßig zur Vermögensbildung oder zum vorsorglichen Sparen verwendet werden können.

¹⁶ Die Differenzierung anhand der aufgewendeten Arbeitszeiten könnte in bestimmten Konstellationen allerdings irreführend sein. So kann ein Unternehmer/eine Unternehmerin, der/die mehrere Gewerbe betreibt, einen hohen Zeitanteil auf ein wirtschaftlich (noch) relativ schwaches Unternehmen konzentrieren (etwa in der Gründungsphase), während ein ertragreiches Unternehmen (etwa wegen eines angestellten Geschäftsführenden) nur wenig Arbeitszeit des/der Selbstständigen beansprucht. Eine Differenzierung anhand alternativer Kennzahlen wie Umsatz oder Ertrag wäre allerdings erhebungspraktisch schwer umzusetzen, weil diese Größen oft nicht oder nur unpräzise messbar sind.

¹⁷ Die zentralen Angaben zum verfügbaren Haushaltseinkommen und zu regelmäßigen finanziellen Verpflichtungen des Haushalts wurden bereits unter die Basismerkmale auf Haushaltsebene in Abschnitt 4.1 gefasst.

Die folgende Übersicht fasst die zur Erfassung empfohlenen Dimensionen mit Bezug auf die persönlichen Einkommensverhältnisse der selbstständig tätigen Zielperson und weitere Einkünfte im Haushalt zusammen.

Einkünfte von selbstständig Tätigen

- Gesamteinkünfte aus selbstständiger Tätigkeit (inkl. Gewinnentnahmen) im letzten Monat, im letzten Kalenderjahr
- Geldwerte Vorteile im letzten Monat (insgesamt, z.B. Firmenwagen, aber auch Sachentnahmen)
- Fester „Unternehmerlohn“ bzw. Geschäftsführergehalt? In welcher Höhe?
- Volatilität der monatlichen Einkünfte groß – klein (Skala)
- Gesamteinkommen aus abhängiger Beschäftigung im letzten Monat, im letzten Kalenderjahr
- Volatilität der monatlichen Einkünfte aus abhängiger Beschäftigung groß – klein (Skala)
- Weitere persönliche Einkünfte aus Mieten/Pachten, Zinsen, Renten, Unterhalt, Arbeitslosengeld? Normale Höhe im Monat insgesamt?

Weitere Einkünfte im Haushalt

- Einkünfte des Partners/der Partnerin aus abhängiger Beschäftigung? selbstständiger Beschäftigung? jeweils normale Höhe im Monat? Einkünfte des Partners/der Partnerin aus anderen Quellen (Mieten/Pachten, Zinsen, Renten, Arbeitslosengeld, Unterhalt)?
- Bezug von Sozialleistungen? Welche (Kindergeld, ALG II, Wohngeld, Elterngeld)?
- Möglichkeit, monatlich einen Betrag zu sparen? Wenn ja, zur Vermögensbildung/zum vorsorglichen Sparen? Wenn ja, jeweils in welcher Höhe?

4.4 Vermögen

Für eine umfassende Bewertung der wirtschaftlichen und sozialen Lage von Selbstständigen sind Daten zu den (Netto-)Vermögensverhältnissen sehr wichtig. Allerdings sind Fragen zu den Vermögen für viele Befragte noch sensibler als solche zu den Einkommensverhältnissen. In der Tendenz ist daher mit einer verminderten Antwortbereitschaft zu rechnen.¹⁸ Darüber hinaus können Antwortausfälle oder Fehleinschätzungen auftreten, weil die Befragten vorhandene Vermögensgegenstände nicht oder nur schwer präzise bewerten können. Die am Markt zu erzielenden Werte im Fall einer Veräußerung sind gerade bei volatilen Märkten schwer zu beurteilen,

¹⁸ Fehlende Vermögenswerte lassen sich mittels multipler Imputationsverfahren ersetzen. Diese haben eine höhere Qualität, wenn Längsschnittdaten zur Verfügung stehen. Jedoch können Imputationsverfahren auch dann Anwendung finden, wenn nicht auf vergangene Werte zurückgegriffen werden kann (Grabka und Westermeier (2015); Schröder et al. (2020)).

(Wieder-)Beschaffungswerte und aktuelle Marktpreise können, etwa bei Bestandsimmobilien, deutlich voneinander abweichen.

Dennoch lassen sich Vermögensdaten im Rahmen bevölkerungsrepräsentativer Befragungen erheben. Dabei werden die Vermögensbestände in der Regel, zum Beispiel in der EVS, auf der Haushaltsebene erfasst. Prinzipiell lässt sich das Vermögen jedoch in einer „persönlichen Vermögensbilanz“ auch individuell erheben. Die methodische Vorgehensweise ist eine Besonderheit des SOEP. Da im vorliegenden Zusammenhang jedoch nur ein einzelnes Haushaltsmitglied – die aus der Bruttostichprobe ausgewählte selbstständig erwerbstätige Person – befragt wird, bietet sich eine Erfassung des Vermögens auf der Haushaltsebene an. Damit werden Fragen zur Aufteilung von Werten auf die Haushaltsmitglieder vermieden, wenn sich Vermögensgegenstände in deren gemeinsamem Besitz befinden.

Die Items zur Erfassung einzelner Vermögensgegenstände können an etablierte Befragungen zur Analyse der Vermögensverteilung in Deutschland – insbesondere SOEP und EVS – angelehnt werden. Im Fall von Eigentümern eines gewerblichen Betriebs (oder Beteiligungen an einem solchen Betrieb) kann beispielsweise als Vermögenswert der Preis vor Steuern, der bei einem Verkauf des Betriebes bzw. der Beteiligung erzielt werden könnte, nach Abzug bestehender Belastungen durch Kredite angesetzt werden. Die untenstehende Übersicht fasst die verschiedenen zu erfassenden Bestandteile des Haushaltsvermögens zusammen. Dabei ist jeweils zu erfassen, ob der Vermögensgegenstand vorhanden ist. Ist dies der Fall, ist dessen Wert von den Befragten zu beziffern. Ist der Wert unbekannt, können dabei gegebenenfalls auch Schätzwerte angegeben werden. Für eine vollständige Vermögensbilanz müssen eventuelle Restschulden aus vorhandenen Krediten ebenfalls erfasst werden.

Bestandteile der Vermögensbilanz auf Haushaltsebene

- Vermögenswert von Betrieben und Betriebsbeteiligungen
- Verkehrswert des (selbst- oder fremdgenutzten) Wohneigentums abzüglich Belastung mit Darlehen
- Geldanlagen: Sparguthaben, Spar- und Pfandbriefe, Aktien, Investmentanteile
- Vertragsguthaben bei Bausparverträgen, Lebensversicherungen und privaten Rentenversicherungen
- Restschuld von Bank-, Privat- oder Ausbildungskrediten (BAföG)

In Analysen zur Vermögensverteilung in Deutschland wird angesichts des bevorstehenden demografischen Übergangs zunehmend die Rolle von Erbschaften und Schenkungen thematisiert (Baresel et al. 2021). Vor diesem Hintergrund kann es zur Beurteilung der wirtschaftlichen Dispositionsfähigkeit von Selbstständigen sinnvoll sein, auch Informationen zu den in einer Vorausschau über mehrere Jahre erwarteten größeren Erbschaften oder Schenkungen zu erheben. Retrospektiv – im Rückblick der letzten 15 Jahre – wurden entsprechende Informationen zum Beispiel mehrfach in den Vermögensmodulen des SOEP erhoben.

4.5 Versicherung und Absicherung gegen zentrale Risiken

4.5.1 Alterseinkommen

Fragen zur sozialen Absicherung der Selbstständigen im Alter sind vor dem Hintergrund andauernder Debatten um die Einbeziehung dieser Gruppe in die Gesetzliche Rentenversicherung bedeutend. Sie sollten sich sowohl auf den aktuellen Erwerb von Rentenanwartschaften beziehen als auch das private Vorsorgeverhalten, inklusive einer etwaigen Absicherung des Lebenspartners oder der Lebenspartnerin, in den Blick nehmen.

Aktueller Erwerb von Rentenanwartschaften

- Mitgliedschaft in einem Versorgungswerk? Künstlersozialkasse? Beitragshöhe, Belastung durch Beiträge (Skala)
- Pflichtversicherung als Selbstständige(r) in der DRV? Art der Pflichtversicherung, Höhe und Art des Beitrags, Belastung durch Beiträge (Skala)
- Freiwillige Versicherung als Selbstständige(r) in der DRV? Höhe des Beitrags, Modus der Beitragsentrichtung, Belastung durch Beiträge (Skala)
- Beiträge zur DRV durch weitere Versicherung, Betriebsrente/Zusatzversorgung bei hybrider Erwerbstätigkeit? Beitragshöhe
- Regelmäßige Einzahlung in private Altersvorsorge? Höhe, Art (Riester, Rürup inkl. Sparplan, LV/Bausparen, ETF, Aktien-, Fondssparplan), Bewertung der Belastung durch private Vorsorge (Skala)

Erwartete Alterseinkommen

- Bestehender Anspruch auf Altersrente der DRV? Erwartete Höhe?
- Bestehender Anspruch auf Betriebsrente? Zusatzversorgung? Pension? Versorgungswerk? Jeweils erwartete Höhe?
- Erwartete Rentenlücke [= Differenz zwischen Einkommen im Ruhestand und letztem Einkommen]
- Qualität des erwarteten Lebensstandards im Alter? (Skala), Vergleich zum aktuellen Lebensstandard
- Auskömmliches Einkommen im Ruhestand erwartet? Bedarf an Grundsicherung im Alter?

Darüber hinaus wären Angaben zu den kumulierten Rentenansprüchen, gegebenenfalls differenziert nach den verschiedenen Säulen des Altersvorsorgesystems, sowie Einschätzungen zur erwarteten Rentenlücke wünschenswert. Allerdings dürften viele Personen im Erwerbsalter zu diesen Variablen keine oder nur unpräzise Zahlenangaben machen können. Ob sich entsprechende Items in der geplanten Befragung umsetzen lassen, müsste daher zunächst in einem sorgfältigen Pretest beurteilt werden. Eine Alternative ist der Rückgriff auf qualitative Abschätzungen, etwa bezüglich des

erwarteten Lebensstandards im Alter oder die Erwartung eines auskömmlichen Einkommens im Ruhestand.

Die obenstehende Übersicht fasst die Dimensionen zusammen, die nach Möglichkeit im Hinblick auf die vorhandene Absicherung von selbstständig Tätigen erfasst werden sollten.

4.5.2 Einkommensausfall

Im Hinblick auf die Versicherung gegen Einkommensausfälle vor dem Eintritt in den Ruhestand ist zwischen eher kurzfristigen und eher längerfristigen Risiken zu unterscheiden. Eher kurzfristige Einkommensausfälle entstehen im Zusammenhang mit dem Risiko des Auftrags- oder Umsatzausfalls. Eher längerfristige Einkommensausfälle sind mit den Risiken der Erwerbsminderung, Berufsunfähigkeit oder Invalidität verbunden. Es ist zu erwarten, dass die zu befragenden Personen relativ genaue Angaben dazu machen können, ob sie über Versicherungen für diese verschiedenen Risiken verfügen. Wie im Zusammenhang der Rentenanwartschaften ist jedoch auch hier damit zu rechnen, dass Angaben zur Höhe der bei Eintritt des Versicherungsfalles zu erwartenden Leistungen nur schwer zu erheben sind. Wiederum könnten alternativ qualitative Einschätzungen erhoben werden, ob der bestehende Versicherungsschutz als ausreichend angesehen wird.

Insbesondere, wenn es um die Absicherung kurzfristiger Einkommensrisiken geht, kann gerade bei nicht sozialversicherungspflichtig Beschäftigten die private Vorsorge über die Bildung von Rücklagen eine wesentliche Rolle spielen. Deshalb sollte zusätzlich erhoben werden, inwieweit die Befragten selbst oder im Haushaltszusammenhang über ausreichende liquide Mittel oder Ersparnisse verfügen, um vorübergehende Ausfälle an mit ihrer selbstständigen Erwerbstätigkeit erzielten Umsätzen oder Einkommen zu überbrücken.

Die nachfolgende Übersicht fasst die verschiedenen Elemente zusammen, die im Zusammenhang mit der privaten Vorsorge oder der Versicherung gegen kürzer- und längerfristige Einkommensausfälle erhoben werden sollten.

Absicherung gegen Einkommensausfälle

- Mitgliedschaft in der ALV: freiwillige Mitgliedschaft auf Antrag aktuell/Monate in den letzten 30 Monaten, Mitgliedschaft als abhängig Beschäftigte(r) aktuell/Monate in den letzten 30 Monaten
- Wie lange reichen vorhandene liquiden Reserven/Ersparnisse aus, um einen Ausfall des eigenen Einkommens aus Selbstständigkeit zu überbrücken? Bis zu drei, sechs, zwölf Monaten, oder länger? Gar nicht? Sind dies eigene Mittel oder Mittel des Haushalts?
- Versicherung(en) gegen Erwerbsminderung, Berufsunfähigkeit, Invalidität vorhanden? Welcher Art? Wird der vorhandene Versicherungsschutz als ausreichend angesehen?

4.5.3 Gesundheit und Pflege

Fragen, auf welche Weise sie kranken- und pflegeversichert sind, sollten von den zu befragenden Zielpersonen im Allgemeinen ohne größere Schwierigkeiten zu beantworten sein. Entsprechende Items sind in anderen bevölkerungsrepräsentativen Befragungen, etwa im SOEP, vorhanden und können als Vorbild dienen. Die im Zusammenhang mit der Art der Versicherung zu erfassenden zentralen Aspekte sind in der nachfolgenden Übersicht zusammengefasst.

Art der Versicherung gegen Gesundheitsrisiken

- Gesetzlich (GKV), nur privat (PKV), gesetzlich plus private Zusatzversicherung
- Wenn GKV: Pflichtmitgliedschaft, mitversichertes Familienmitglied, freiwillig als Selbstständiger
- Wenn PKV: Wer ist versichert? Monatlicher Beitrag? Selbstbeteiligung? ggf. welcher Art

Bewertung der vorhandenen Versicherung

- Zufriedenheit mit Leistung und Kosten der Versicherung (Skala)
- Falls ein anderer Versicherungsstatus rechtlich möglich wäre, wäre dieser gewünscht?
- Beiträge sehr hohe Belastung oder kein Problem (Skala)
- Wenn PKV: Wechsel in Basistarif/in GKV vollzogen oder angestrebt

Weil Erwerbstätige bei der Wahl der Art ihrer Kranken- und Pflegeversicherung bzw. ihres Versicherungsstatus durch gesetzliche Vorgaben eingeschränkt sind, müssen die beobachteten Versicherungsverhältnisse nicht den von den befragten Personen gewünschten entsprechen. Daher empfiehlt es sich, zusätzlich Items in die geplante Befragung aufzunehmen, die auf die Bewertung der vorhandenen Versicherung gegen Gesundheitsrisiken abzielen. Konkret könnte die Zufriedenheit mit der Leistung (auch im Verhältnis zu den Beiträgen), oder der Wunsch nach einem anderen Versicherungsstatus, wenn dieser rechtlich zulässig wäre, abgefragt werden. Darüber hinaus könnten mit der Versicherung verbundene wirtschaftliche Risiken erkannt werden, indem eine Einschätzung dazu erhoben wird, wie stark die finanzielle Belastung durch die Versicherungsbeiträge empfunden wird, und ob bei Privatversicherten ein Wechsel in den günstigeren Basistarif oder in die GKV angestrebt wird.

4.6 Erwerbsbiografie

Die Charakterisierung der Erwerbssituation der selbstständig tätigen Zielpersonen ist, wie in Abschnitt 4.2 ausgeführt, am aktuellen zeitlichen Rand bzw. dem Erwerbsverlauf in einer relativ kurzen Frist vor dem Befragungszeitpunkt (bevorzugt: zwölf Monate) verankert. Als wesentliche Hintergrundinformation sollte aber auch der erwerbsbiografische Verlauf davor – und damit potenzielle Risikofaktoren bei der Entwicklung hin zu einer aktuell prekären Selbstständigkeit – erfasst werden. Im Idealfall würde die Erwerbsbiografie entweder mittels eines Kalendariums („Wie war der Erwerbsstatus in einem bestimmten Jahr? Wie im Jahr darauf? etc.“) oder einer sequenziellen Abfrage („Wann war die erste Beschäftigungsepisode? Was folgte als nächste Episode? etc.“) im Detail abgefragt. Dies ist allerdings ein sehr aufwändiges – und befragte Personen leicht ermüdendes

oder überforderndes – Verfahren. Dies gilt insbesondere, wenn dabei auch der Möglichkeit Rechnung getragen werden soll, dass Personen mehreren – selbstständigen oder nicht-selbstständigen – Tätigkeiten parallel nachgehen. Deshalb erscheint eine summarische Erfassung wesentlicher erwerbsbiografischer Informationen angebracht, um die Wirtschaftlichkeit der Befragung zu gewährleisten und das Risiko von Interviewabbrüchen zu begrenzen.

Summarische Erfassung der Erwerbsbiografie

- Jahr der Aufnahme der ersten Beschäftigung nach Abschluss der primären beruflichen Ausbildung
- Jahr der Aufnahme der ersten selbstständigen Beschäftigung nach Abschluss der primären Ausbildung
- Gesamtzahl der Jahre seit Abschluss der primären Ausbildung in Beschäftigung (insgesamt, sozialversicherungspflichtig, selbstständig, in Teilzeit), Gesamtzahl der Jahre ohne Beschäftigung (davon nicht erwerbstätig wegen Arbeitslosigkeit, wegen familiärer Verpflichtungen)
- Dauer des längsten Beschäftigungsverhältnisses mit einem Arbeitgeber? Anzahl der bisherigen Arbeitgeber insgesamt?
- Dauer der längsten Episode ununterbrochener selbstständiger Beschäftigung
- Wechsel zwischen abhängiger Beschäftigung bei einem Arbeitgeber und selbstständiger Tätigkeit. Offene Abfrage der Anzahl oder kategoriale Abfrage der Häufigkeit
- Schon einmal parallel abhängig und selbstständig beschäftigt gewesen? Wenn ja, Häufigkeit hybrider Episoden? Wie viele Jahre hybride Beschäftigung insgesamt?

Die obenstehende Übersicht veranschaulicht, wie die erwerbsbiografische Entwicklung einer zu befragenden Person mit relativ wenigen Items summarisch erfasst werden könnte. Zu erheben wären dafür insbesondere der Zeitpunkt ihres Eintritts in den Arbeitsmarkt nach Beendigung der Ausbildung und der Zeitpunkt der Aufnahme der ersten selbstständigen Tätigkeit. Darüber hinaus wären die in verschiedenen Formen der Erwerbstätigkeit (inklusive hybrider Beschäftigung) jeweils insgesamt verbrachte Zeit, sowie die kumulierten Zeiten von Erwerbsunterbrechungen (insbesondere wegen Arbeitslosigkeit und familiärer Verpflichtungen) zu erfassen. Schließlich könnten Indikatoren für die Stabilität des erwerbsbiografischen Verlaufs erfasst werden, wie zum Beispiel die Frequenz des Arbeitgeberwechsels bzw. die durchschnittliche Beschäftigungsdauer bei einem Arbeitgeber, oder die Häufigkeit der Wechsel zwischen abhängiger und selbstständiger Beschäftigung.

4.7 Subjektive Einschätzungen, Motivationen und Haltungen sowie Persönlichkeitsmerkmale und Zufriedenheit

An die geplante Datenerhebung lassen sich weitere Fragenblöcke andocken, die für eine genauere Erfassung von für den Erfolg am Arbeitsmarkt oder die Vermeidung prekärer sozialer Lagen relevanten individuellen Eigenschaften (wie Persönlichkeitsmerkmale bzw. „soft skills“) oder für spezielle Analysen, etwa zur Bedeutung subjektiven Einschätzungen, Motivationen und Haltungen, genutzt werden können. Die Möglichkeit der Aufnahme entsprechender Fragenblöcke dürfte die Attraktivität des vom BMAS geplanten Vorhabens für Auftragnehmer mit wissenschaftlicher Ausrichtung stärken. Die Erweiterung des Fragenprogramms um eher „weiche“ Aspekte im Fragenprogramm bringt aber auch erhebungstechnische Vorteile. Der Kontrast zu den übrigen bisher behandelten, überwiegend faktendominierten Themenkomplexen erlaubt eine lebendigere Fragebogen-Dramaturgie. So können Interviewabbrüche wegen Ermüdung oder erlahmender Motivation der Beteiligten reduziert werden. Die folgende Übersicht zeigt beispielhaft, wie das Fragenprogramm um die Dimensionen der individuellen Einstellungen bezüglich der gesetzlichen Sozialversicherungen und einer möglichen eigenständigen sozialen Versicherung für Selbstständige, der für unternehmerisches Handeln besonders relevanten Persönlichkeitsmerkmale, sowie der empfundenen Arbeitsqualität und Arbeitsbelastungen erweitert werden könnte.

Einstellungen und Persönlichkeitsmerkmale

- Haltung zu Sozialversicherungen (differenziert GRV, GKV, ALV): Akzeptanz allgemein, Bereitschaft zu einer freiwilligen Mitgliedschaft, Haltung zu einer Pflichtmitgliedschaft, akzeptierter Monatsbeitrag, finanziell verkraftbare Mindest-/Höchstbeiträge
- Haltung zu eigenständigen sozialen Versicherungen für Selbstständige (differenziert nach Risiken Alter, Gesundheit, Erwerbseinkommensausfälle), gewünschte/akzeptierte Leistungsparameter (Leistungsumfang, Selbstbehalte, fixe vs. variable Beiträge)
- Persönlichkeitsmerkmale: Unternehmergeist, Risikobereitschaft, Kontrollüberzeugung

Empfundene Arbeitsqualität und Arbeitsbelastungen

- Gesundheit und Stress, Zukunftsängste
- Qualität des vorhandenen Arbeitsschutzes
- Zufriedenheit: Allgemeine Lebenszufriedenheit, Allgemeine Arbeitszufriedenheit, Zufriedenheit mit dem Einkommen, mit der Arbeitszeit

Grundsätzlich wären auch weitere (oder andere) inhaltliche Blöcke denkbar, um Erkenntnisinteressen des BMAS zu bedienen. So könnten spezifische Fragenkomplexe für besondere Teilzielgruppen – etwa Selbstständige in der Gründungsphase oder mittels digitaler Plattformen selbstständig Tätige – gestaltet werden. Entsprechende Vertiefungsfragen sind allerdings nur dann sinnvoll, wenn dieser Bevölkerungsanteil der fokussierten Teilstichprobe hinreichend groß ist oder die realisierte Stichprobe so geschichtet wird, dass die für belastbare zielgruppenspezifische Aussagen erforderlichen Beobachtungszahlen erreicht werden.

Andere denkbare inhaltliche Vertiefungen, die an die gesamte Zielpopulation zu richten wären, könnten sich auf besondere (Prekaritäts-)Erfahrungen im Verlauf der Corona-Krise oder der Ukraine-Krise beziehen. Allerdings ist zu empfehlen, das geplante Vorhaben, das primär auf die Schließung elementarer Wissenslücken im Hinblick auf die Risiken prekärer sozialer Lagen von selbstständig Beschäftigten und den sie strukturell begünstigenden Risikofaktoren abzielt, nicht mit an aktuellen Ereignissen und kurzfristigen Erkenntnisinteressen ausgerichteten Elementen zu überfrachten. Das in den vorangegangenen Abschnitten umrissene Erhebungsprogramm ist bereits reichlich umfangreich, der Spielraum für die Integration zusätzlicher Themenfelder somit eng. Damit die Erhebung praktisch durchführbar bleibt, müssten zum Ausgleich wohl Items aus dem Kernfragenprogramm gestrichen werden. Die Alternative wäre ein Fragebogen-Split, bei dem Zielpersonen zufällig jeweils nur ein Teil des Fragenprogramms zugewiesen wird. Dadurch sinkt jedoch die realisierte Fallzahl pro Item. Dies muss entweder durch eine Vergrößerung der Nettostichprobe ausgeglichen werden, oder die Belastbarkeit der erzielten Befunde sinkt bei unveränderter Größe der Nettostichprobe.

5. Durchführung des Vorhabens

5.1 Einleitung

In der Gesamtbetrachtung ergibt sich aus den in den beiden vorangegangenen Kapiteln ausgeführten Überlegungen, dass die Durchführung einer bevölkerungsrepräsentativen Erhebung für eine Studie zur sozialen Lage von Selbstständigen grundsätzlich machbar ist. Dafür gibt es zwei verschiedene Ansatzpunkte: zum einen kommerzielle Unternehmensdatenbanken, welche die Grundgesamtheit der wirtschaftsaktiven Unternehmen in Deutschland so gut wie vollständig abbilden, zum anderen bevölkerungsrepräsentative Zufallsstichproben. Bei beiden Ausgangsbasen ist allerdings zunächst ein Screening erforderlich, um die Zielpopulation der geplanten Untersuchung, das heißt Personen mit selbstständigen Tätigkeiten, zu identifizieren und daraus eine Bruttostichprobe für eine umfassende Befragung zu generieren. Des Weiteren existieren in anderen bevölkerungsrepräsentativen Umfragen bereits erprobte Items, anhand derer sich die einzelnen Bestandteile des multidimensionalen Konzepts „soziale Lage“ für Erwerbstätige operationalisieren lassen. Diese eignen sich größtenteils auch für Erwerbstätige, die – ganz oder teilweise – selbstständig arbeiten. Im Rahmen der geplanten Untersuchung wäre primär eine Reihe von Items zu ergänzen, die sich auf Besonderheiten der sozialversicherungsrechtlichen Absicherung und der von Selbstständigen in der Regel als Substitut zu betreibenden privaten Risikoversorge beziehen.

Allerdings ist der zu erwartende Aufwand für die Durchführung einer repräsentativen Studie zur sozialen Lage von Selbstständigen beträchtlich. Dies liegt erstens am eher niedrigen Anteil der Selbstständigen an der erwerbstätigen Bevölkerung (gemessen für die Haupterwerbstätigkeit etwa zehn Prozent) bei gleichzeitig starker Heterogenität der Erscheinungsformen selbstständiger Arbeit. Um spezifische repräsentative Aussagen für wesentliche Teilpopulationen ableiten zu können, muss eine große Nettostichprobe realisiert werden. Dies wiederum erfordert ein umfangreiches Screening zur Gewinnung einer großen Bruttostichprobe, weil der Status der Selbstständigkeit in den zur Verfügung stehenden Auswahlgesamtheiten nicht direkt beobachtbar ist. Die zu realisierenden Stichproben sollten dabei möglichst im Hinblick auf zu untersuchende spezifische Arten der Selbstständigkeit geschichtet werden, um hinreichende Beobachtungszahlen zu erhalten. Zweitens erfordert eine hochwertige, multidimensionale Erfassung sozialer Lagen und der darauf einwirkenden zentralen Faktoren den Einsatz eines umfangreichen Fragebogens. Deshalb ist mit relativ hohen Erhebungskosten pro Interview zu rechnen. Die zu erreichende hohe Fallzahl in Kombination mit dem zu erwartenden hohen Aufwand je Fall bringt es mit sich, dass für die Umsetzung der beabsichtigten Studie nicht nur ein substanzielles Budget, sondern auch ein ausreichend dimensionierter zeitlicher Rahmen eingeplant werden muss.

Die Grundfestlegung, die geplante Untersuchung bevölkerungsrepräsentativ anzulegen, hat zur Folge, dass für Gruppen mit selten anzutreffenden Konstellationen der Selbstständigkeit, bei denen wegen fehlender Merkmale in den Auswahlgesamtheiten auch keine Schichtung möglich ist, keine belastbaren Ergebnisse erreicht werden. Um solche Teilpopulationen wirtschaftlich zu erfassen, müsste spezifisches Adressmaterial erschlossen und ein gesonderter Feldzugang gewählt werden, wobei wahrscheinlich Abstriche bei der Repräsentativität hinzunehmen wären. Entsprechende spezifische Studien ließen sich später an das hier erörterte Vorhaben andocken, indem dafür das gleiche Erhebungsinstrument verwendet wird. So ließe sich konsistent bewerten, inwiefern die sozialen Lagen der gesondert zu fokussierenden Teilgruppen vom repräsentativen Durchschnitt der Bevölkerung mit selbstständigen Tätigkeiten abweichen.

Wie in Abschnitt 2 erörtert, existieren bereits bevölkerungsrepräsentative Datenquellen, die für Analysen zu den wirtschaftlichen und sozialen Verhältnissen von Selbstständigen geeignet sind, und

die für diesen Zweck auch genutzt werden. Dies gilt insbesondere für den Mikrozensus und das SOEP, die zudem jährlich durchgeführt werden, so dass sich damit die Verhältnisse nah am aktuellen zeitlichen Rand sowie Veränderungen über die Zeit – im Querschnitt und beim SOEP als Panel-Studie auch im individuellen Längsschnitt – untersuchen lassen. Dennoch bräuchte eine eigenständige Erhebung zur sozialen Lage von Selbstständigen, wie sie vom BMAS geplant ist, in mehrfacher Hinsicht einen substanziellen Mehrwert:

- Es lässt sich ein umfangreiches und speziell auf den Untersuchungszweck fokussiertes Fragenprogramm realisieren, wie es in Abschnitt 4 umrissen wurde. Die Integration eines solchen Programms in eine generelle amtliche Erhebung wie den Mikrozensus wäre in der Gesamtheit nicht, in sehr kleinen Teilen nur mit erheblichem zeitlichen Vorlauf möglich. Das SOEP ist im Hinblick auf die Integration neuer Fragen und Themenkomplexe zwar flexibler; zudem wäre ein Teil der notwendigen Items schon im regulären Kern der Erhebung enthalten. Auch hier ergeben sich jedoch materielle und organisatorische Beschränkungen, weil der Fragebogen nicht beliebig verlängert werden kann, somit die Aufnahme neuer Themenschwerpunkte in Konkurrenz zu anderen Ideen für thematische Erweiterungen steht, und redaktionelle Entscheidungen über die Ausrichtung des Kernfragebogens einige Vorlaufzeit in Anspruch nehmen. Prinzipiell steht für Untersuchungen mit sehr spezifischen Erkenntnisinteressen auch die SOEP-Innovations-Stichprobe (SOEP-IS) zur Verfügung. Der Katalog an aktuell nicht im SOEP enthaltenen Items für die vom BMAS geplante Untersuchung ist jedoch für diesen Rahmen zu umfangreich; zudem würden damit zu geringe Fallzahlen erreicht.
- Es können hinreichend große Fallzahlen realisiert werden, um relevante Teilpopulationen von selbstständig Erwerbstätigen differenziert zu analysieren. Da der Mikrozensus jeweils rund ein Prozent der Bevölkerung in Deutschland zu ihren Lebens- und Arbeitsbedingungen befragt, ist auch eine große Anzahl von Selbstständigen enthalten, so dass Heterogenitäten in dieser Zielpopulation prinzipiell gut untersucht werden können. Bei dieser Datenbasis sind die eingeschränkten Möglichkeiten, hybride Erwerbskonstellationen zu erfassen und fehlende oder zu grob klassifizierte Variablen zur Operationalisierung prekärer sozialer Lagen von Selbstständigen die zentralen limitierenden Faktoren. Beim SOEP stößt man bei einer Anzahl von 1.000 Selbstständigen, die derzeit in der regulären Erhebung enthalten sind, bei tiefer gegliederten Auswertungen rasch an Grenzen. Prinzipiell besteht zwar die Möglichkeit, in das SOEP hinreichend große spezifische Stichproben von Personen bzw. Haushalten zu integrieren, die eine bestimmte Teilbevölkerung repräsentieren. Eine unabhängig konzipierte Befragung dieser Teilbevölkerung kann jedoch flexibler, insbesondere ohne Rücksicht auf besondere Anforderungen durch die (potenziell spätere) Überführung in eine dem SOEP inhärente Panel-Struktur, angelegt und organisiert werden.

Im Folgenden werden die bei der Durchführung einer bevölkerungsrepräsentativen Erhebung zur sozialen Lage von Selbstständigen vorzunehmenden zentralen Festlegungen weiter konkretisiert und der sich daraus ergebende Finanz- und Zeitbedarf grob abgeschätzt. Die Stichprobenbildung und die Erhebungsverfahren im Feld betreffend, werden dabei verschiedene Alternativen gegenübergestellt, um eine Bandbreite – in einem späteren Vergabeverfahren zu dem vom BMAS geplanten Vorhaben durch die Bieter zu konkretisierender – möglicher Herangehensweisen aufzuzeigen.

5.2 Erhebungsmodus

Das in Kapitel 4 skizzierte Fragenprogramm umfasst zur Operationalisierung der Zielgröße „soziale Lage“, bzw. „prekäre Selbstständigkeit“ überwiegend Elemente, die bei allen Zielpersonen gleichermaßen zu erheben sind. Allerdings können in manchen Fällen – insbesondere bei hybriden Arbeitsverhältnissen – Filterführungen erforderlich sein, um etwa sozialversicherungsrechtlich komplexen Konstellationen gerecht zu werden. Alles in allem ist im Rahmen der geplanten Erhebung jedoch nur mit verhältnismäßig wenig komplexen Filterführungen zu rechnen – vorausgesetzt, erwerbsbiografische Informationen werden nur summarisch abgefragt. Von dem Versuch, den beruflichen Karriereweg in einer zeitlichen Abfolge zu erfassen, ist abzuraten. Der Fragebogen würde durch eine detaillierte Abfrage erwerbsbiografischer Verläufe überdehnt, da schon die Erfassung der verschiedenen Dimensionen der sozialen Lage den Einsatz umfangreicher Items erfordert.

Sollen die in Kapitel 4 umrissenen Befragungsinhalte angemessen umgesetzt werden, ist von einer durchschnittlichen Interviewdauer von mindestens 35 bis 40 Minuten auszugehen. Bei dieser Fragebogenlänge scheiden selbstadministrierte strukturierte Interviews per CAWI- oder PAPI-Modus aus. Da eine Kombination von selbstadministrierten und mündlich-persönlichen Verfahren oder ein Split der Fragebogen zur Verkürzung der Befragungsdauer zu einer schlechteren Datenqualität führen kann, sollte die Erhebung entweder im CATI- oder CAPI-Modus durchgeführt werden.¹⁹ Weil das Erhebungsprogramm einige sensible Inhalte (Details zu Einkommen, Vermögen, Risikoversorge) umfasst, bringt dies auch den Vorteil, dass qualifizierte Interviewerinnen und Interviewer Antwortverweigerungen durch geschickte Kommunikation entgegenarbeiten können. Sie können darüber hinaus Nachfragen zu einigen nicht zu vermeidenden komplexen Begrifflichkeiten und Abgrenzungen beantworten.

Der Anteil der Selbstständigen mit ausländischer Staatsangehörigkeit an allen Selbstständigen ist annähernd genauso hoch wie der Anteil der Personen mit ausländischer Staatsangehörigkeit an allen Erwerbstätigen (zuletzt: rund zwölf Prozent). Dabei sind in der Teilgruppe der Solo-Selbstständigen Personen mit ausländischer Staatsangehörigkeit leicht überproportional vertreten, in der Gruppe der Selbstständigen mit abhängig Beschäftigten leicht unterproportional (Bonin et al. 2022). In einer repräsentativen Stichprobe ist der zu erwartende Anteil von Personen mit ausländischer Staatsbürgerschaft demnach recht klein. Angesichts dessen ist zu erwarten, dass sich potenzielle Verzerrungen, weil mangels Deutschsprachkenntnissen Interviews nicht (vollständig) geführt werden können, in Grenzen halten. Somit erscheint die Erstellung fremdsprachiger Erhebungsinstrumente nicht wirtschaftlich. Zur Begrenzung von (Antwort-)Ausfällen könnten alternativ bei Zielpersonen, die schlecht Deutsch können, mehrsprachige Interviewerinnen oder Interviewer eingesetzt werden, die bei auftretenden Verständnisproblemen punktuell Hilfe durch Übersetzen geben können.

5.3 Abgrenzung der Zielgruppe

Eine Grundfestlegung für die geplante Studie ist, selbstständige Arbeit möglichst breit zu erfassen. Insbesondere soll sich die Untersuchung nicht auf Selbstständigkeit im Haupterwerb beschränken. Des Weiteren sollen in der Bevölkerung nur selten vorkommende Formen selbstständiger Tätigkeiten nicht von vornherein ausgeschlossen werden. Auch wenn damit eine umfassende Abdeckung der Zielgruppe angestrebt ist, erscheinen aus erhebungspraktischen Gründen einige grundlegende Abgrenzungen sinnvoll.

¹⁹ Für eine ausführliche Diskussion der Vor- und Nachteile der verschiedenen Erhebungsmodi und kombinierter Befragungsweisen siehe oben Abschnitt 3.3.

Zunächst empfiehlt sich eine Beschränkung der Untersuchung auf die Bevölkerung bzw. die Erwerbstätigen im Haupterwerbsalter – konkret auf Personen im Alter von 18 bis 67 Jahren. Hierfür spricht, dass sich bei Selbstständigkeit von Jugendlichen (die etwa in den Bereichen Gaming, Influencing oder IT tendenziell zunimmt) oder selbstständiger Tätigkeit von Rentnerinnen und Rentnern die Absicherung zentraler wirtschaftlicher Risiken häufig grundlegend anders darstellen dürfte als bei Personen im Haupterwerbsalter. Die Einbeziehung dieser besonderen Teilpopulationen würde daher Anpassungen bzw. spezifische Erweiterungen des Erhebungsinstruments erfordern; zugleich wäre der Aufwand, eine an den Altersrändern ausreichende Bruttostichprobe zu generieren, wegen der zu erwartenden niedrigen Inzidenz selbstständiger Tätigkeiten beträchtlich.

Grundsätzlich könnte die Abgrenzung des Status „selbstständig“ anhand von formal-rechtlichen Kriterien oder Selbstauskünften vorgenommen werden. Erhebungspraktisch betrachtet kommt jedoch eigentlich nur das Prinzip der Selbstauskunft in Frage, wie es in aller Regel auch in anderen Befragungen von Erwerbstätigen, etwa im Rahmen des Mikrozensus, umgesetzt wird. Beim Versuch einer formal-rechtlichen Abgrenzung müssten die Befragten in bestimmten Konstellationen nicht nur umfangreiche Erläuterungen erhalten. Auch würde das Licht auf rechtliche Dunkelzonen geworfen, wie etwa Scheinselbstständigkeit oder informelle selbstständige Tätigkeiten. Beides könnte erhebliche Unit-Nonresponse nach sich ziehen. Entscheidet man sich daher für eine Abgrenzung der Zielgruppe anhand von Selbstauskünften, sind im Ergebnis gewisse Unschärfen in Bereichen unvermeidlich, für die Statusunsicherheit charakteristisch ist oder in denen die rechtliche Bewertung, ob es sich um eine selbstständige Tätigkeit oder abhängige Beschäftigung handelt, von den genauen Umständen im Einzelfall abhängt. Dies betrifft zum Beispiel die Plattformarbeit in der Gig Economy, Geschäftsführerinnen und Geschäftsführer, die zugleich Anteile an einer Gesellschaft halten, die Bewertung der Vermietung von (Wohn-)Eigentum oder des Betriebens einer Photovoltaikanlage²⁰ als selbstständige Tätigkeit.

Ein anders gelagertes Abgrenzungsproblem ergibt sich im Hinblick auf selbstständige Arbeit, die nur unregelmäßig geleistet wird. Erfasst man den Erwerbsstatus zu einem Stichtag – etwa dem Datum oder dem Monat vor der Befragung – werden zwar gelegentliche, aber trotzdem wirtschaftlich bedeutsame selbstständige Tätigkeiten systematisch untererfasst. Zugleich kann die gemessene Inzidenz der Selbstständigkeit saisonale Schwankungen reflektieren.²¹ Wählt man eine zeitraumbezogene Klassifikation, zum Beispiel anhand der Erwerbsverhältnisse in den vorangegangenen zwölf Monaten, werden auch einmalig oder selten ausgeübte selbstständige Tätigkeiten erfasst, die für die Lebensumstände und die soziale Lage der Person eigentlich nicht prägend sind. Um die Spannweite zwischen diesen Polen abstecken zu können, bietet es sich an, sowohl ein stichtags- als auch ein periodenbezogenes Messkonzept zu verwenden.

Inwieweit es sich bei den so erfassten selbstständigen Tätigkeiten um marginale Arbeit handelt, die für die Lebensumstände bzw. soziale Lage nicht prägend ist, muss anhand zusätzlicher Kriterien operationalisiert werden. Das primäre Kriterium dafür wäre der absolute zeitliche Umfang der selbstständigen Tätigkeit in Stunden. Der Cut-off sollte dabei allerdings eher niedrig angesetzt werden, damit nicht zu viele hybride Arbeitsverhältnisse mit selbstständigen Tätigkeiten im Nebenerwerb, an denen ein besonderes Erkenntnisinteresse besteht, herausfallen. Als weiteres Abgrenzungskriterium kommt auch das mit der selbstständigen Tätigkeit erzielte Einkommen bzw.

²⁰ Wer Strom aus einer privaten Photovoltaikanlage ausschließlich mit Verkaufs- und Gewinnerzielungsabsicht in das öffentliche Netz einspeist, zählt steuerrechtlich als Unternehmer.

²¹ Dieses Problem lässt sich bei einer hinreichend langen Feldphase begrenzen. Werden etwa die Interviews über einen Zeitraum von zwölf Monaten verteilt, lässt sich der Einfluss der Saisonalität herausrechnen, indem über die Monatswerte der stichtagsbezogenen Inzidenzen selbstständiger Tätigkeiten gemittelt wird.

der Anteil des erzielten Einkommens am gesamten Erwerbseinkommen in Frage. Dieser Ansatz birgt allerdings das Risiko, dass prekäre selbstständige Tätigkeitsverhältnisse, die durch ein geringes Einkommen pro Stunde charakterisiert sind, systematisch unberücksichtigt bleiben. Deshalb sollte ein Einkommens-Cut-off nur in Kombination mit einem Stunden-Cut-off verwendet werden.

Es empfiehlt sich, erwerbstätige Personen, die nur einen geringfügigen Teil ihrer Arbeitszeit selbstständig tätig sind und mit dieser Tätigkeit nur einen geringfügigen Teil ihres gesamten Erwerbseinkommens erzielen, aus der Studie herauszunehmen, um die Ergebnisse nicht durch die wirtschaftlichen Verhältnisse nicht prägende marginale Selbstständigkeit zu verzerren. Dies ist auch deshalb erforderlich, weil die geplante Nettostichprobe mit Rücksicht auf die Wirtschaftlichkeit der Untersuchung nicht beliebig vergrößert werden kann und auch genügend Fälle mit substanzieller Selbstständigkeit zur Auswertung vorhanden sein müssen. Wie häufig das vorgeschlagene Abgrenzungskriterium in der Kombination aus Stunden- und Einkommens-Cut-off erfüllt ist, lässt sich ex ante nicht bestimmen, sondern kann erst im Rahmen der Untersuchung anhand der einlaufenden Befragungsergebnisse festgestellt werden. Erweisen sich die Cut-off-Punkte als zu niedrig, läuft also ein zu hoher Anteil von Personen in die Nettostichprobe, müssen diese im Verlauf eventuell nach oben korrigiert werden, damit der zur Auswertung verfügbare Pool an selbstständig Tätigen eine angemessen ausgewogene Struktur erhält.

5.4 Umfang und Strukturierung der Nettostichprobe

Die vom BMAS geplante Untersuchung soll differenzierte Aussagen zu den Risikofaktoren bei selbstständig Tätigen mit prekären Verhältnissen ermöglichen, sich jedoch nicht auf diese Teilgruppe beschränken, sondern die Verteilung der finanziellen und sozialen Lagen von Selbstständigen möglichst umfassend abbilden. Zugleich ist das Ziel, belastbare differenzierte Aussagen für nach sozio-demografischen Merkmalen, insbesondere aber auch für nach Typen selbstständiger Erwerbsverhältnisse gegliederte Teilgruppen zu erhalten.

Eine fundamentale Schwierigkeit, eine diesen Zielgrößen gerecht werdende Nettostichprobe zu planen, liegt darin, dass die Verteilung wesentlicher Merkmale in der Zielpopulation der Selbstständigen nicht bekannt ist. Die hängt vor allem damit zusammen, dass im vorliegenden Zusammenhang, wie in Abschnitt 5.3 erörtert, eine weite Abgrenzung der Zielgröße angestrebt wird, die insbesondere auch eher geringfügig und hybrid selbstständig Tätige einschließt. Diese Zielpopulation unterscheidet sich von der Abgrenzung gängiger Statistiken zur Selbstständigkeit, die somit ungeeignet sind, die zu erwartenden Verteilungen der für disaggregierte Auswertungen heranzuziehenden Merkmale in der relevanten Grundgesamtheit verlässlich abzuschätzen. Ebenfalls fehlen geeignete Daten, um vorab zu beurteilen, welcher Anteil der Zielpopulation sich in einer prekären sozialen Lage befindet. Um weiterführende Analysen bezüglich der Determinanten von Prekarität bei selbstständig Erwerbstätigen vornehmen zu können, muss in der realisierten Nettostichprobe jedoch eine hinreichende Zahl von Beobachtungen vorhanden sein. Welche Zahl hinreichend ist, um systematische Risikofaktoren statistisch zu identifizieren, hängt wiederum von deren – ex ante ebenfalls nur schwer einzuschätzenden – Varianz ab.

Angesichts dieser Einschränkungen lässt sich der Umfang einer für die Ziele der geplanten Studie adäquaten Nettostichprobe nur heuristisch bestimmen. Erfahrungsgemäß sollten in einer Merkmalsdimension mindestens 500 bis 1.000 Beobachtungen vorhanden sein, um Repräsentativität zu beanspruchen und darauf aufbauende, tiefer gegliederte multivariate Auswertungen vornehmen zu können. Angesichts dessen wird empfohlen, eine Nettostichprobe mit mindestens 5.000 Fällen zu realisieren:

- Unterstellt man, dass gut jeder siebte selbstständig Erwerbstätige in der Grundgesamtheit prekär beschäftigt ist,²² würden bei einer reinen Zufallsauswahl etwa 750 Fälle in der Nettostichprobe vorkommen; durch moderate Schichtung etwa bezüglich des Merkmals der Solo-Selbstständigkeit, bei der prekäre soziale Lagen häufiger zu vermuten sind als bei Gewerbetreibenden mit Beschäftigten, ließe sich eine noch günstigere Fallzahl erreichen.
- Angenommen, dass gut zwei Prozent der arbeitenden Bevölkerung in Deutschland hybrid selbstständig sind, und unterstellt, dass jeder zweite hybrid Selbstständige im Nebenerwerb selbstständig ist, würden bei einer reinen Zufallsauswahl aus der Grundgesamtheit etwa 1.000 Fälle in die Nettostichprobe gelangen.²³
- Geht man davon aus, dass ein Prozent der Beschäftigten in Deutschland Plattformarbeit betreibt und dabei Selbstständigen-Status hat, würden bei einer reinen Zufallsauswahl gut 500 Fälle in die Nettostichprobe gelangen.²⁴ Damit wären Auswertungen für die Gesamtpopulation von Plattform-Beschäftigten machbar, bei einer Differenzierung nach verschiedenen Arten der Plattformarbeit (Gigwork vs. Crowdwork) stieße man dagegen an Fallzahlengrenzen.

Diese Beispiele veranschaulichen, dass mit dem empfohlenen Umfang der Nettostichprobe bereits ein in wesentlichen Dimensionen differenziertes belastbares Bild gezeichnet kann, das auch noch Spielraum lässt für Auswertungen bezüglich kombinierter Merkmale (etwa Gliederung der hybrid selbstständig Erwerbstätigen nach Geschlecht und Alter) sowie für informative multivariate Schätzungen (etwa zu den Merkmalen, die prekäre Selbstständigkeit begünstigen).

Die Möglichkeiten, zu differenzierten disaggregierten Auswertungen zu gelangen, lassen sich noch verbessern, indem die Auswahlgesamtheit geschichtet wird, so dass Personen mit eher selten auftretenden Merkmalen mit höherer Wahrscheinlichkeit in die Nettostichprobe gelangen. Um repräsentative Aussagen zu erhalten, müssen die anhand einer geschichteten Stichprobe gewonnenen Ergebnisse allerdings durch Anwendung von Hochrechnungsfaktoren adjustiert werden. Dafür müssen die Randverteilungen der Schichtungsmerkmale in der Grundgesamtheit bekannt sein; zugleich muss darauf geachtet werden, dass einzelne inverse Auswahlwahrscheinlichkeiten nicht zu groß werden, um Verzerrungen durch einzelne mit einem hohen Gewicht versehene Ausreißer zu vermeiden.

Bei einer Strukturierung der Stichprobe durch Schichtung sollten zumindest die zentralen Kategorien von Selbstständigen – konkret gewerbetreibende Selbstständige mit abhängig Beschäftigten (eventuell grob kategorisiert nach Anzahl von Beschäftigten), gewerbetreibende Selbstständige ohne abhängig Beschäftigte bzw. Solo-Selbstständige sowie Angehörige der Freien Berufe – gleichmäßig beachtet werden. Vor dem Hintergrund der Diskussion um die soziale Lage und Absicherung von

²² Gemäß empirischen Untersuchungen von Schulze Buschoff et al. (2017) waren 2014 in Deutschland 17 Prozent der Solo-Selbstständigen prekär beschäftigt; nimmt man Selbstständige mit Beschäftigten hinzu, dürfte der Anteil niedriger liegen.

²³ Gruber (2019) stellt fest, dass gemäß Mikrozensus im Jahr 2015 2,1 Prozent der Erwerbstätigen mindestens zwei Erwerbstätigkeiten ausüben, wovon mindestens eine selbstständig ausgeübt wurde. Der Anteil der Selbstständigen an allen Erwerbstätigen lag 2020 bei 8,9 Prozent (Bonin et al. 2022). Bei dieser Quote sind allerdings nur Erwerbstätige berücksichtigt, die ausschließlich oder bei hybrider Erwerbstätigkeit im Haupterwerb selbstständig tätig sind.

²⁴ Die amtliche Statistik liefert bislang nur wenige Erkenntnisse zur Verbreitung von Plattformarbeit. Die Inzidenz von Plattformarbeit wurde häufig im Rahmen von Online-Befragungen gemessen, die in der Regel in diesem Kontext unplausibel hohe Werte generieren. Eine telefonisch basierte Umfrage durch Bonin und Rinne (2017) ermittelte einen Anteil von 0,9 Prozent Erwerbstätigen mit Plattformarbeit.

selbstständig Tätigen sollten jedoch auch weitere Schichtungsdimensionen ins Auge gefasst werden, die einander nicht wechselseitig ausschließen müssen, wie insbesondere:

- hybrid selbstständige Erwerbstätigkeit,
- Selbstständigkeit in Teilzeit,
- Selbstständige mit Tätigkeiten im Dienstleistungsbereich vs. Selbstständige in anderen Wirtschaftsbereichen,

Spezifischere sozialpolitisch relevante Teilpopulationen dagegen dürften für eine gezielte Schichtung in einer bevölkerungsrepräsentativen Ausgangsgesamtheit entweder zu schwach besetzt sein oder im Rahmen eines Screenings zur Stichprobenbildung nicht mit wirtschaftlich vertretbarem Aufwand hinreichend exakt abzugrenzen sein. Hierzu zählen etwa Gründerinnen und Gründer,²⁵ Selbstständige mit Plattform-Arbeit, selbstständige Helferinnen und Helfer, insbesondere in der Landwirtschaft, oder selbstständige Kunstschaffende.

5.5 Feldzugang

5.5.1 Einleitung

Die zu erreichende Stichprobe muss von der Personenebene und nicht von der Unternehmensebene ausgehend strukturiert werden, weil dieselbe selbstständig tätige Person gleichzeitig mit mehreren Unternehmen in Bezug stehen kann. Zudem steht ein Teil der selbstständig tätigen Personen in keinem Bezug zu einem registrierten Unternehmen; dies betrifft primär die Angehörigen der Freien Berufe. Daher sollte eine bevölkerungsrepräsentative Zufallsstichprobe als Anker der geplanten Befragung gewählt werden.

Um ausgehend von einer bevölkerungsrepräsentativen Stichprobe, in der die gesuchte Eigenschaft der selbstständigen Erwerbstätigkeit nicht direkt erfasst ist, eine Bruttostichprobe von Personen zu gewinnen, die der Zielgruppe der Untersuchung angehören, ist im ersten Schritt ein Screening erforderlich. Wie in Abschnitt 3.2.2 ausführlich erörtert, bietet sich hierfür als effiziente Möglichkeit die Nutzung einer laufenden bevölkerungsrepräsentativ angelegten Omnibus-Befragung an, vorzugsweise im CATI-Modus.²⁶ Dies hat den Vorteil, dass automatisch eine Reihe von sozio-demografischen Merkmalen miterfasst wird, die genutzt werden können,

- um Selektivitäten hinsichtlich der Teilnahmebereitschaft an der Kernbefragung zu erkennen; hierzu müssen die Verteilungen der sozio-demografischen Merkmale von Selbstständigen in der Ausgangsgesamtheit der in der Omnibus-Befragung erfassten Selbstständigen, und der Selbstständigen in der Bruttostichprobe (das heißt, der Selbstständigen, die ihre Bereitschaft zur Teilnahme an der Kernbefragung erklärt haben) miteinander verglichen werden;

²⁵ Hier besteht die Schwierigkeit, dass die Neugründung eines Unternehmens nicht identisch mit der Tatsache ist, dass eine Person eine selbstständige Tätigkeit begonnen hat. Tatsächlich werden viele Neugründungen durch Selbstständige vollzogen, die bereits als Gewerbetreibende tätig sind. Demnach wäre das zu bevorzugende Abgrenzungskriterium, ob in jüngerer Zeit eine neue selbstständige Tätigkeit begonnen wurde. Hieraus lässt sich allerdings nicht schließen, dass dahinter die Neugründung eines Unternehmens steht, weil dieses Kriterium auch bei einer Unternehmensübernahme erfüllt ist. Für eine fokussierte Untersuchung zur sozialen Lage von Gründerinnen und Gründern erscheint ein spezifischerer Feldzugang, etwa anknüpfend an den KfW-Gründungsmonitor, sinnvoll.

²⁶ Ein CAPI-Modus wäre wegen der Kürze des Screening-Interviews unwirtschaftlich. Im CAWI-Modus wären der auf der Screening-Stufe wünschenswerte hohe Ausschöpfungsgrad und die Repräsentativität wahrscheinlich deutlich schwerer zu erreichen. Diese Argumente sprechen auch gegen eigenständig durchgeführte Screening-Umfragen auf Grundlage bevölkerungsrepräsentativen Adressmaterials im PAPI-Modus oder im gemischten PAPI/CAWI-Modus.

- um die effektiven Auswahlwahrscheinlichkeiten der an der Kernbefragung Teilnehmenden, das heißt, der Selbstständigen in der Nettostichprobe, mit Bezug auf bestimmte sozio-demografische Merkmale zu bestimmen; dies ist die Basis für die Hochrechnung der Auswertungen der Nettostichprobe auf die in der Omnibus-Befragung erfasste Grundgesamtheit der Selbstständigen.

Allerdings gelangen beim Screening einer bevölkerungsrepräsentativen Zufallsstichprobe kleine Teilpopulationen mit nur geringen Fallzahlen in die für die anschließende Kernbefragung zu verwendende Bruttostichprobe. In dieser wären somit insbesondere die Gewerbetreibenden mit einer größeren Anzahl von Beschäftigten, die einen kleinen Anteil an der Grundgesamtheit ausmachen, nur relativ schwach abgebildet. Da eine Weiterführung des Screenings, bis ausreichende Fallzahlen auch bei dieser Zielgruppe vorhanden sind, bei zu dünn besetzten Teilzielgruppen keine wirtschaftliche Option darstellt, werden im Folgenden verschiedene Möglichkeiten erörtert, mit diesem Problem umzugehen.

5.5.2 Aufbau einer dualen Stichprobe

Ein Ansatzpunkt ist der Aufbau einer dualen Stichprobe, bei dem Kontaktdaten aus einer umfassenden Unternehmensdatenbank komplementär für eine gezielte Erweiterung der aus der bevölkerungsrepräsentativen Zufallsstichprobe zu gewinnenden Bruttostichprobe genutzt werden. Die Wahrscheinlichkeit, über diesen Zugangskanal auf gewerbetreibende Selbstständige zu treffen, die sich zur Aufnahme in die Bruttostichprobe eignen, ist substantiell höher als beim Screening im Rahmen der zufälligen Bevölkerungstichprobe. Darüber hinaus bietet sich die Möglichkeit, die auf diesem Weg in die Bruttostichprobe gelangenden gewerbetreibenden Selbstständigen anhand weiterer in den Unternehmensdatenbanken vorhandener Merkmale zu schichten. Diese Merkmale könnten zum Beispiel der Sektor oder die Anzahl der Beschäftigten der Unternehmen sein, mit denen eine Zielperson in Bezug steht.

Beim Screening der Selbstständigen auf Basis von Adressmaterial aus einer Unternehmensdatenbank sollte aus Konsistenzgründen möglichst das gleiche Befragungsverfahren genutzt werden wie beim Screening im auf der bevölkerungsrepräsentativen Zufallsstichprobe basierenden Befragungsstrang. Allerdings stehen in den Unternehmensbanken teilweise keine Telefonnummern der Zielpersonen zur Verfügung, so dass ein Screening im CATI-Modus nicht durchweg ohne Weiteres umsetzbar ist. Da Zielpersonen mit fehlenden Telefonnummern ohnehin schriftlich über vorliegende E-Mail- oder postalische Adressen kontaktiert werden müssen,²⁷ bietet sich hier ein selbstadministrierter Erhebungsmodus an – wegen der Kosteneffizienz bevorzugt im CAWI-Modus. Dieses Verfahren kommt in Frage, da in Unternehmensdatenbanken eine hohe Anzahl an Bruttofällen zur Verfügung steht. Die im CAWI-Modus niedrige Ausschöpfungsquote auf der Screening-Ebene kann somit durch einen Einsatz von mehr Adressen ausgeglichen werden. Die Unternehmensdatenbanken liefern zudem Strukturinformationen für die nahezu vollständige Grundgesamtheit der Selbstständigen, die in Verbindung mit einem amtlich registrierten Unternehmen stehen. Diese Informationen können genutzt werden, um durch den CAWI-Modus bedingte Selektivitäten zu erkennen und gegebenenfalls mittels Gewichtungsfaktoren zu korrigieren.

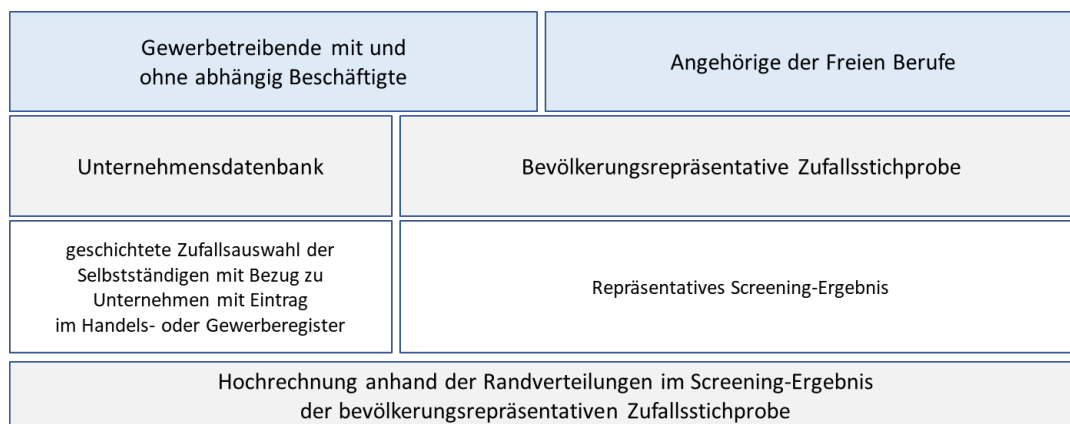
Beim Aufbau einer dualen Bruttostichprobe dient das Screening nicht nur dazu zu klären, ob die angesprochene Person zur Zielgruppe gehört und zur Teilnahme an der Kernbefragung bereit ist. Zusätzlich müssen grundlegende sozio-demografische Hintergrundvariablen erfasst werden, wie sie im Rahmen von Omnibus-Befragungen (und somit für die anhand der bevölkerungsrepräsentativen

²⁷ Eine kontaktlose Recherche der Telefonnummern wäre mit erheblichem Aufwand verbunden und dürfte nur lückenhaft möglich sein.

Zufallsstichprobe identifizierten Selbstständigen) routinemäßig erhoben werden, aber in Unternehmensdatenbanken nicht oder nur unvollständig vorhanden sind. Indem diese Merkmale ebenfalls für die ausgehend von einer Unternehmensdatenbank rekrutierten Zielpersonen erfasst werden, lassen sich die Daten der Kernerhebung für diese Gruppe anhand der Randverteilungen der sozio-demografischen Merkmale für Gewerbetreibende mit und ohne abhängig Beschäftigte in der bevölkerungsrepräsentativen Zufallsstichprobe hochrechnen. Sie können damit konsistent mit den Ergebnissen der Kernbefragung für die im von der bevölkerungsrepräsentativen Zufallsstichprobe ausgehenden Erhebungsstrang erfassten Gewerbetreibenden zusammengeführt werden. Eine konsistente Zusammenführung setzt ferner voraus, dass für die Kernerhebung dasselbe Befragungsverfahren verwendet wird, wie in Abschnitt 5.2 erörtert, wegen des Umfangs und der Komplexität des Erhebungsprogramms also ein mündlich-persönlicher Modus mit Einsatz von Interviewpersonal. Dafür müssen beim Screening im CAWI-Modus gegebenenfalls fehlende geeignete Kontaktdaten erhoben werden.

Mit Blick auf die Zusammenführung ist beim Aufbau einer dualen Stichprobe kein völliger Split in dem Sinn zu empfehlen, dass Gewerbetreibende nur über eine Unternehmensdatenbank identifiziert werden und Angehörige der Freien Berufe oder sonstige selbstständig Erwerbstätige ohne angemeldetes Gewerbe über eine bevölkerungsrepräsentative Zufallsstichprobe. Um die in zwei unterschiedlich angelegten Befragungssträngen für die gleiche Teilpopulation erzielten Ergebnisse der Kernbefragung internen Konsistenzchecks unterziehen zu können, muss auch im auf der bevölkerungsrepräsentativen Zufallsstichprobe basierenden Erhebungsstrang eine ausreichende Anzahl von Gewerbetreibenden erfasst und befragt werden.

Abbildung 5.1 Aufbau einer dualen Stichprobe



Quelle: Darstellung des IZA.

Abbildung 5.1 fasst das Vorgehen beim Aufbau einer dualen Stichprobe schematisch zusammen. Es gibt hierbei zwei Erhebungsstränge, von denen einer auf die Angehörigen der Freien Berufe und andere selbstständig Erwerbstätige ohne angemeldetes Gewerbe fokussiert, der andere auf die Gewerbetreibenden mit und ohne abhängig Beschäftigte, deren Grundgesamtheit in verfügbaren kommerziellen Unternehmensdatenbank praktisch vollständig abgebildet ist. Diese Zielgruppe wird allerdings auch im von einer bevölkerungsrepräsentativen Zufallsstichprobe ausgehenden Erhebungsstrang miterfasst. In diesem Strang wird per Screening eine Bruttostichprobe von Selbstständigen ohne und mit angemeldetem Gewerbe aufgebaut, deren Struktur (wenn keine planmäßige Schichtung vorgenommen würde) die Verteilung der verschiedenen Arten von Selbstständigkeit in der Ausgangsgesamtheit der Bevölkerung in Deutschland abbildet. Die auf dieser

Basis erzielten Befragungsergebnisse in der sich anschließenden Kernerhebung sind als repräsentativ zu betrachten.

Im auf einer Unternehmensdatenbank basierenden Erhebungsstrang werden ebenfalls Selbstständige – nicht Unternehmen – als Untersuchungseinheiten betrachtet. Personen, die mit mehreren als Gewerbe angemeldeten Unternehmen verbunden sind, werden also nur einmal erfasst. Die komplementäre Nutzung der Unternehmensdatenbank für den Feldzugang dient dem Ziel, solche Gruppen von Selbstständigen anzusprechen, die im auf Repräsentativität hin angelegten anderen Erhebungsstrang wegen ihres geringen Anteils an der Population der Selbstständigen nur selten in die Befragung eingehen; dies gilt insbesondere für die wegen der Dominanz von Kleinbetrieben in der Unternehmenslandschaft relativ kleine Anzahl an Selbstständigen, die mit mittleren und großen Betrieben verbunden sind. Diese werden bei der für das im ersten Schritt vorzunehmende Screening zu treffenden Zufallsauswahl aus dem Adressmaterial der Unternehmensdatenbank bewusst bevorzugt. Dies ist möglich, weil die Datenbank Hintergrundinformationen zu den mit der Person verbundenen Unternehmen umfasst. Somit könnten auch Selbstständige, die mit Unternehmen aus bestimmten Sektoren der Wirtschaft verbunden sind, gezielt in die Befragung hineingesteuert werden. Auf dieser Auswahlebene geplante Verzerrungen lassen sich später durch Gewichtungsfaktoren korrigieren, weil die Unternehmensdatenbank praktisch das Universum der zum Gewerbe angemeldeten Unternehmen abbildet. Zielpersonen, die die im Screening abgefragten Bedingungen erfüllen und für eine weitere Befragung zur Verfügung stehen, werden danach im gleichen Modus befragt wie die im anderen Erhebungsstrang adressierten Zielpersonen.

Um die Resultate der Kernbefragungen in beiden Erhebungssträngen schließlich zusammenzuführen, ist eine Hochrechnung auf die Grundgesamtheit der Erwerbstätigen in Deutschland vorzunehmen. Da diese nur innerhalb der bevölkerungsrepräsentativen Zufallsstichprobe abgebildet ist, nicht jedoch in den auf Gewerbeanmeldungen basierenden Unternehmensdatenbanken, sind in beiden Strängen Hochrechnungsfaktoren einzusetzen, die sich anhand der Randverteilungen sozio-demografischer Merkmale der im Screening-Prozess erfassten Zielpersonen (mit und ohne Bereitschaft zur Teilnahme an der Folgebefragung) im bevölkerungsrepräsentativen Strang bilden lassen.

5.5.3 Alternativen

Der Aufbau einer dualen Stichprobe, wie im vorangegangenen Abschnitt skizziert, erfordert einen erheblichen Aufwand. Es müssen dafür zwei verschiedene Datenbasen erschlossen werden, was jeweils unterschiedliche Kompetenzen erfordert. Die Umbasierung von der unternehmensbezogenen Struktur der Unternehmensdatenbanken auf eine personenbezogene Struktur, wie sie für das geplante Vorhaben benötigt ist, ist nicht trivial und birgt größere Fehlerrisiken. Die Qualität der vorliegenden Kontaktdaten zu den Zielpersonen und der Anteil der tatsächlich selbstständig im Sinne des Vorhabens tätigen Kontaktpersonen ist vorab schwer zu beurteilen. Ein weiteres Risiko liegt im zumindest teilweise erforderlichen Wechsel zum CAWI-Modus für das Screening von Zielpersonen mit lediglich digitalen Adressdaten im auf einer Unternehmensdatenbank basierenden Strang. Neben einer hohen Rate von Nicht-Teilnahmen bereits auf der Screening-Stufe ist zusätzlich beim Wechsel zu einem mündlich-persönlichen Erhebungsmodus für die Kernbefragung mit Reibungsverlusten zu rechnen. Schließlich lassen sich durch die Ergänzung des zweiten Strangs zwar höhere Fallzahlen für die Teilpopulation der Gewerbetreibenden mit einer größeren Anzahl an Beschäftigten generieren; ein limitierender Faktor bleibt aber, dass die zur Hochrechnung der erzielten Ergebnisse verfügbaren Faktoren auf einer relativ geringen Fallzahl von Gewerbetreibenden im bevölkerungsrepräsentativen Strang basieren und entsprechend mit größerer Unsicherheit behaftet sind.

Da zudem die Inzidenz prekärer sozialer Lagen bei Selbstständigen, die mit größeren Unternehmen verbunden sind, relativ gering sein dürfte und diese Teilzielgruppe entsprechend weniger im Fokus

sozialpolitischer Debatten steht, sollten auch mögliche Alternativen zum Aufbau einer dualen Stichprobe in Betracht gezogen werden. Konkret könnten auch folgende Schritte die Umsetzung des Vorhabens vereinfachen, ohne die zentralen Erkenntnisinteressen des BMAS zu vernachlässigen:²⁸

- die Anlage einer einsträngigen Erhebung, die nur auf einer bevölkerungsrepräsentativen Erhebung basiert. Allerdings sollte hierbei, um dennoch eine hinreichende Zahl von Interviews mit mittleren oder größeren Unternehmen zu realisieren, zum Ausgleich die per Screening zu gewinnende Bruttostichprobe vergrößert werden. Als zusätzliche Maßnahme könnte die Ansprache der Zielgruppe im Anschluss an das Screening-Interview intensiviert werden, um eine möglichst hohe Ausschöpfung bei den vorhandenen Kontakten zu erreichen. Bei einer Ausweitung des Screenings verlängert sich allerdings die Feldzeit in Abhängigkeit von der Anzahl der im Rahmen der Omnibus-Befragung pro Lauf durchschnittlich zu gewinnenden Anzahl an Bruttofällen.
- die Anlage einer Erhebung auf Basis einer bevölkerungsrepräsentativen Erhebung, die jedoch um eine auf Grundlage einer Unternehmensbedank gewonnenen, zufällig gezogenen Stichprobe von Gewerbetreibenden ergänzt wird, welche über eine größere Zahl von Beschäftigten (etwa: zehn und mehr) verfügen und für die eine verwertbare Telefonnummer vorliegt. Sofern das Fehlen einer Telefonnummer in den Unternehmensbanken nicht mit der sozialen Lage der Zielpersonen korreliert, könnten die mit einer solchen Aufstockung der Bruttostichprobe verbundenen Verzerrungen begrenzt sein.²⁹ Zur Hochrechnung der für diese zusätzliche Teilstichprobe erzielten Ergebnisse müssen wie beim Aufbau einer dualen Stichprobe die anhand des Screenings der bevölkerungsrepräsentativen Zufallsstichprobe gewonnenen Randverteilungen sozio-demografischer Merkmale verwendet werden.

5.6 Kostenabschätzungen

Welche Kosten mit der Durchführung des Vorhabens gemäß den verschiedenen oben entwickelten Alternativen verbunden sind, lässt sich nur schwer abzuschätzen. Dies liegt zum einen daran, dass die Kalkulationsgrundlagen möglicher Anbieter nicht öffentlich sind, und zum anderen enthält das Mengengerüst der Untersuchung eine Reihe von schwer abzuschätzenden Parametern. Die hier vorgenommenen Überschlagsrechnungen liefern daher nur grobe Anhaltspunkte des zu erwartenden Kostenaufwands. Mit Bezug auf die Kernerhebung wird durchweg unterstellt, dass diese im CATI-Modus durchgeführt wird und die durchschnittliche Dauer pro Interview bei 35 Minuten liegt. Bei einer Durchführung im CAPI-Modus würden substantiell höhere Kosten pro Fall entstehen.

²⁸ Eine weitere denkbare Alternative wäre, die vorgegebene Zielpopulation zu beschränken, etwa durch eine grundsätzliche Herausnahme von Gewerbetreibenden mit mehr als einer bestimmten Anzahl von Beschäftigten (etwa: zehn und mehr). Allerdings ginge damit der Anspruch verloren, ein vollständiges Bild der sozialen Lage von Selbstständigen in Deutschland zu erhalten, und die Studienergebnisse dürften die Inzidenz prekärer Selbstständigkeit eher überzeichnen, weil eine Teilpopulation mit mutmaßlich niedriger Inzidenz bewusst nicht repräsentiert wäre. Hierdurch würde die Studie stark angreifbar, weshalb diese Möglichkeit im Folgenden nicht weiter betrachtet wird.

²⁹ Um diese Annahme zu plausibilisieren, müsste geprüft werden, inwieweit sich die einer Unternehmensdatenbank zu entnehmenden personen- und unternehmensbezogenen Daten systematisch unterscheiden, je nachdem, ob für die über diesen Kanal identifizierten Selbstständigen eine Telefonnummer vorhanden ist oder nicht. Ist dies der Fall, müssten anhand der beobachteten Variablen Korrekturfaktoren bestimmt werden, um Verzerrungen infolge nicht zufällig fehlender Kontaktdaten zu vermeiden.

5.6.1 Aufbau einer dualen Stichprobe

Die überschlägige Schätzung des mit der Durchführung der Studie mittels Aufbau einer dualen Stichprobe verbundenen Aufwands geht vom in Abschnitt 5.4 begründeten Ziel einer Nettostichprobe von 5.000 Personen aus, bei denen die Kernbefragung erfolgreich durchgeführt wird. Zudem wird vorgegeben, dass davon eine Mehrheit (60 Prozent) ausgehend von einer bevölkerungsrepräsentativen Zufallsstichprobe rekrutiert wird. Bei den übrigen 2.000 Nettofällen soll es um basierend auf einer Unternehmensdatenbank rekrutierte Gewerbetreibende handeln.

Um 3.000 Nettofälle aus der zufälligen Bevölkerungsgesamtheit heraus zu rekrutieren, müssen überschlägig gerechnet gut 110.000 Erwachsene gescreent werden. Die Erwerbstätigenquote in dieser Population liegt bei etwa 45 Prozent, so dass zunächst eine Auswahlgesamtheit von rund 50.000 erwerbstätigen Personen erreicht wird. Von diesen ist allerdings nur ein Bruchteil selbstständig erwerbstätig. Die aktuelle Rate der ausschließlich oder im Haupterwerb selbstständig Erwerbstätigen an allen Erwerbstätigen erreicht derzeit etwa zehn Prozent. Da im Rahmen der geplanten Untersuchung jedoch auch im Nebenerwerb selbstständig Erwerbstätige erfasst werden sollen, wird diese Quote ad hoc auf zwölf Prozent erhöht. Hieraus ergibt sich eine erwartete Gesamtzahl von 6.000 Selbstständigen, von denen, wenn man die Struktur der Selbstständigen im Haupterwerb zugrunde legt, gut über die Hälfte Solo-Selbstständige oder Angehörige der Freien Berufe sind. Es ist jedoch davon auszugehen, dass nicht alle im Screening erfassten Selbstständigen bereit sind, an der späteren Kernbefragung teilzunehmen. Unterstellt man, dass 60 Prozent der Selbstständigen ohne Beschäftigte und 40 Prozent der Selbstständigen mit Beschäftigten teilnahmebereit sind,³⁰ liefert dieser Strang der Erhebung im Endeffekt für die erste Teilgruppe etwa 2.000 Fälle, für die zweite Teilgruppe etwa 1.000 Fälle.

Das unterstellte Mengengerüst für den auf einer Unternehmensdatenbank basierenden zweiten Erhebungsstrang stellt sich wie folgt dar. Als Zielgrößen wurden je 1.000 Nettofälle vorgegeben, die per CATI und per CAWI gescreent und rekrutiert werden.³¹ Um diese Fallzahlen zu erreichen, muss zunächst eine deutlich höhere Bruttostichprobe angesprochen werden, weil mit einer niedrigen Ausschöpfungsrate zu rechnen ist. Dabei kumulieren sich potenziell mehrere Faktoren: (1) die in der Datenbank enthaltenen Kontaktdaten sind wahrscheinlich mit Fehlern behaftet, so dass ein Teil der Kontaktversuche ins Leere läuft; (2) ein Teil der kontaktierbaren Personen gehört nicht zur Zielgruppe, sondern arbeitet tatsächlich als angestellte Führungskraft; (3) die Bereitschaft zur Teilnahme an der Hauptbefragung ist tendenziell niedrig, insbesondere in der Gruppe der Gewerbetreibenden mit einer größeren Zahl an Beschäftigten, die in diesem Erhebungsstrang bevorzugt angesprochen werden sollen. Bei der im CAWI-Modus anzusprechenden Teilgruppe tritt zu diesen Faktoren noch die erfahrungsgemäß niedrige Antwortrate bei selbstadministrierten Erhebungen hinzu. Ausgehend von diesen Überlegungen wird (konservativ) unterstellt, dass das

³⁰ Die unterstellte Differenz lässt sich damit begründen, dass bei den Unternehmerinnen und Unternehmern mit Beschäftigten eine größere Aversion gegen Datenerfassungen zu vermuten ist, und dass Selbstständige mit größerem Geschäft wegen eines dichtereren Terminplans und längeren Arbeitszeiten schwerer zu erreichen sind. Die unterstellten Quoten der Teilnahmebereitschaft sind recht hoch angesetzt, weil die Zielpersonen durch ihre Mitwirkung an der Omnibus-Befragung bereits ein allgemeines Interesse an Umfragen gezeigt haben. Dass die Kernerhebung auf Themen fokussiert, zu denen sie einen engen persönlichen Bezug haben, kann dieses Interesse verstärken. Die Chancen, über Omnibus-Befragungen Erwerbstätige für spezifische Folgebefragungen zu interessanten Themen zu gewinnen, sind erfahrungsgemäß recht gut. Allerdings lässt sich schwer einschätzen, inwieweit sich dies auf die spezielle Zielgruppe der Selbstständigen übertragen lässt. Sollten sich zu niedrige Quoten der Teilnahmebereitschaft ergeben, müssten eventuell Incentivierungen eingesetzt werden.

³¹ Da im CAWI-Modus das Risiko von systematischen Antwortausfällen größer ist, sollte der Anteil nicht zu hoch ausfallen. Andererseits spricht die größere Effizienz für eine möglichst große Zahl von per CAWI rekrutierten Fällen. Die vorgegebene Gleichverteilung der CATI- und CAWI-Fälle soll zudem gewährleisten, dass Selektivitäten, die sich in unterschiedlichen Randverteilungen bei den beobachteten sozio-demografischen Merkmalen niederschlagen, recht gut erkannt werden können.

Verhältnis von eingesetzter Brutto- zu Nettostichprobe im CATI-Teil bei 10:1 liegt, im – dafür kosteneffizienten – CAWI-Teil sogar bei 25:1.

Aus diesem Mengengerüst lässt sich, ausgehend von den jeweiligen Fallzahlen und dem jeweiligen Fragenprogramm, der Aufwand für die verschiedenen durchzuführenden Erhebungen anhand von Erfahrungswerten zumindest grob abschätzen. Es wird davon ausgegangen, dass für die CATI-Teile ein professionelles Befragungsunternehmen eingeschaltet wird, während den CAWI-Teil technisches Personal an einer Forschungseinrichtung übernimmt. Weiterhin wird unterstellt, dass ein kostenfreier Zugang zu einer Unternehmensdatenbank für Forschungszwecke besteht.

Bezüglich der einzusetzenden Personalressourcen wird angenommen, dass sämtliche Fragebögen durch nicht gewinnorientiert arbeitende Forscherinnen und Forscher konzipiert werden. Zudem wird davon ausgegangen, dass dieses wissenschaftliche Personal auch die Auswertungen und inhaltlich orientierten Analysen der Daten sowie die Erstellung des Abschlussberichts zur Präsentation der Ergebnisse übernimmt. Für diese Leistungen werden Ressourcen im Umfang von zwei Personenmonaten für eine langjährig erfahrene Projektleitung, vier Personenmonaten für Senior Researcher und sechs Personenmonaten für Junior Researcher angesetzt.

Unter diesen Voraussetzungen ergibt sich ein Gesamtaufwand von rund 560.000 Euro (netto). Dieser setzt sich zusammen wie in Tabelle 5.1 dargestellt. Der Großteil der zu erwartenden Kosten entfällt mit rund 420.000 Euro auf die verschiedenen Erhebungsteile. Das Screening zur Rekrutierung der Teilnehmerinnen und Teilnehmer für die Haupterhebung hat mit 120.000 Euro einen Anteil von gut 28 Prozent an den gesamten Erhebungskosten. Der angesetzte Aufwand für das wissenschaftliche Personal trägt etwa ein Viertel zu den Gesamtkosten bei. Wird die Fragebogenentwicklung und Datenauswertung durch ein kommerzielles Befragungsinstitut übernommen, ist von substantiell höheren Personalkosten auszugehen.

Tabelle 5.1 Geschätzter Kostenaufwand (netto) – Aufbau einer dualen Stichprobe

Kostenposition	Betrag in Euro
Erhebungen, davon	420.000
Screening Bevölkerungsstichprobe (Omnibus)	65.000
Screening Unternehmensdatenbank (50% CAWI/50% CATI)	55.000
Haupt-Befragung (CATI, N = 5.000 Fälle, 35 Minuten)	300.000
Wissenschaftliches Personal	140.000
Gesamtkosten	560.000

Quelle: Berechnungen und Darstellung des IZA

Bezogen auf die unserer groben Kalkulation zugrundeliegende Nettofallzahl von 5.000 beläuft sich der reine Erhebungsaufwand – ohne die Durchführung von Analysen und die Erstellung von Berichten – je realisiertem Hauptinterview auf 84 Euro. Zum Vergleich: Die Realisierung eines vollständigen Interviews im Rahmen des SOEP ist mit etwa 300 Euro zu veranschlagen, Interviews mit stark verkürztem Fragenprogramm mit mindestens 100 Euro.

5.6.2 Alternativen

Wird auf die Nutzung des Feldzugangs über eine Unternehmensdatenbank ganz oder weitgehend verzichtet, verändert sich das Kostentableau spürbar. Für die Kostenabschätzung bezüglich der ersten Alternative, dass die Befragung nur auf einer bevölkerungsrepräsentativen Zufallsstichprobe basiert wird und Fallzahlenprobleme mit Bezug auf Selbstständige mit größerem Gewerbe durch eine Vergrößerung der zu realisierenden Nettostichprobe angegangen werden, wird unterstellt, dass in der Kernerhebung 6.000 statt 5.000 Fälle realisiert werden sollen. Zusätzlich wird – optimistisch – unterstellt, dass die Bereitschaft, an der Kernerhebung teilzunehmen, durch besonders intensive Ansprache und Incentivierungen bei Selbstständigen mit Beschäftigten auf das Niveau bei den Solo-Selbstständigen gesteigert werden kann.³²

Um unter diesen Voraussetzungen 6.000 Nettofälle aus der zufälligen Bevölkerungsgesamtheit heraus zu rekrutieren, müssen überschlägig gerechnet 185.000 Erwachsene gescreent werden. Darunter befinden sich erwartungsgemäß knapp 10.000 Selbstständige. Bei einer unterstellten effektiven Teilnahmequote von 60 Prozent bei beiden Gruppen von Selbstständigen umfasst die realisierte Nettostichprobe schließlich rund 3.300 Solo-Selbstständige und rund 2.700 Selbstständige mit Beschäftigten.

Tabelle 5.2 Geschätzter Kostenaufwand (netto) – Bevölkerungsrepräsentative Erhebung mit um 20 Prozent vergrößerter Nettostichprobe

Kostenposition	Betrag in Euro
Erhebungen, davon	465.000
Screening Bevölkerungsstichprobe (Omnibus)	115.000
Haupt-Befragung (CATI, N = 6.000 Fälle, 35 Minuten)	350.000
Wissenschaftliches Personal	95.000
Gesamtkosten	560.000

Quelle: Berechnungen und Darstellung des IZA

Wie Tabelle 5.2 zeigt, bleibt der Gesamtaufwand im Vergleich der Variante mit Aufbau einer dualen Stichprobe mit 560.000 Euro (netto) gleich. Dabei steigen die Erhebungskosten sogar um rund 45.000 Euro oder gut ein Zehntel. Dieser Anstieg speist sich aus der größeren Nettostichprobe und dem zusätzlichen Aufwand zur Steigerung der Teilnahmebereitschaft bei den Selbstständigen mit Beschäftigten. Dennoch reduziert sich der reine Erhebungsaufwand auf rund 78 Euro pro realisiertem

³² Inwieweit intensivierte Ansprache und Incentivierung die Mitwirkungsbereitschaft bei der speziellen Gruppe erhöhen können, ist a priori nur schwer einzuschätzen. Die Alternative wäre, die Feldzeit für das Screening der repräsentativen Bevölkerungsstichprobe zu verlängern, um die Bruttostichprobe soweit zu vergrößern, dass auch bei einer niedrigen Teilnahmebereitschaft der Selbstständigen mit Beschäftigten diese Personengruppe hinreichend in der realisierten Nettostichprobe vertreten ist. Bei dieser Vorgehensweise würden Mehrkosten für das Screening anfallen, die bei einem sparsamen Design (vgl. Abbildung 3.1) allerdings recht moderat sind; andererseits könnten die Kosten für eine intensivierte Ansprache und Incentivierung entfallen. Wesentlicher limitierender Faktor ist somit der vom BMAS vorgegebene Zeitrahmen der Studie. Ist dieser hinreichend lang, lässt sich das Risiko, dass eine intensivierte Ansprache und Incentivierung zur Befragungsteilnahme nicht hinreichend wirken, durch eine Verlängerung der Screening-Feldzeit vermeiden. Lässt der Zeitrahmen diese Verlängerung nicht zu, und soll wegen der unsicheren Erfolgsaussichten auf eine intensivierte Ansprache und Incentivierung der Selbstständigen mit Beschäftigten verzichtet werden, könnte schließlich auch nur eine Nettofallzahl von 5.000 vorgegeben werden, wie in den übrigen hier kalkulierten Varianten. Da darunter die Belastbarkeit der Befunde mit Bezug auf die Gruppe der Selbstständigen mit Beschäftigten leidet, kann diese Vorgehensweise – je nach den Schwerpunkten des Erkenntnisinteresses des BMAS – jedoch im Vergleich zu den Alternativen weniger wirtschaftlich sein.

Interview. Kompensierend wirkt, dass das aufwändige Screening der im Vergleichsfall von über eine Unternehmensdatenbank zu gewinnenden Selbstständigen entfällt. Darüber hinaus entfallen Personalkosten im Zusammenhang mit der Erschließung der Unternehmensdatenbank, der Prüfung von Selektivitäten und der konsistenten Zusammenführung von Daten aus unterschiedlichen Erhebungssträngen.

Für die Kostenabschätzung bezüglich der zweiten Alternative, dass die Befragung ganz überwiegend auf einer bevölkerungsrepräsentativen Zufallsstichprobe basiert wird und Fallzahlenprobleme durch Ergänzung einer Zufallsstichprobe mit Telefonnummern von Gewerbetreibenden mit mehr als einer bestimmten Anzahl von Beschäftigten angegangen werden, die aus einer Unternehmensdatenbank gezogen wird, wird unterstellt, dass von den 5.000 Selbstständigen in der Nettostichprobe 750 im CATI-Modus über die Unternehmensdatenbank rekrutiert werden.

Wie Tabelle 5.3 zeigt, sind bei dieser Durchführungsvariante sowohl die Gesamtkosten mit (netto) 530.000 Euro als auch die reinen Erhebungskosten mit 410.000 Euro am niedrigsten. Die reinen Erhebungskosten pro realisiertem Kerninterview belaufen sich auf 82 Euro. Im Vergleich zu den beiden anderen durchgerechneten Varianten bewegen sich die Kosten für das wissenschaftliche Personal in der Mitte. Einerseits ist der Personalaufwand gegenüber der zuvor betrachteten Variante höher, weil mit der Erschließung der Unternehmensdatenbank, der Stichprobenziehung daraus und der konsistenten Hochrechnung ein Mehraufwand verbunden ist. Im Vergleich zur einer Anlage der Untersuchung mit einer voll entwickelten dualen Stichprobe ist der Aufwand jedoch geringer, weil weiter ausdifferenzierte Schichtungsrouitinen und die technische Umsetzung des CAWI-Screening entfallen.

Tabelle 5.3 Geschätzter Kostenaufwand (netto) – Bevölkerungsrepräsentative Erhebung mit Ergänzungsstichprobe aus Unternehmensdatenbank

Kostenposition	Betrag in Euro
Erhebungen, davon	410.000
Screening Bevölkerungsstichprobe (Omnibus)	89.000
Screening Unternehmensdatenbank (100% CATI)	21.000
Haupt-Befragung (CATI, N = 5.000 Fälle, 35 Minuten)	300.000
Wissenschaftliches Personal	120.000
Gesamtkosten	530.000

Quelle: Berechnungen und Darstellung des IZA

Im Vergleich der Durchführungsmöglichkeiten erscheint die zuletzt betrachtete Variante besonders attraktiv, weil sie (1) Vorteile des Feldzugangs über eine bevölkerungsrepräsentative Zufallsstichprobe mit denen des Feldzugangs über eine Unternehmensdatenbank miteinander verbindet, (2) planbar eine ausreichende Zahl an Fällen mit substanzieller Selbstständigkeit generiert, (3) dennoch eine relativ einfache Struktur aufweist, (4) ohne einen mit Risiken behafteten Modus-Wechsel bei den Befragungen auskommt, und (5) einen – wenn auch relativ geringen – Kostenvorteil aufweist.

Der Vergleich der verschiedenen Varianten verdeutlicht auch, dass die Befragungskosten stets den weitaus größten Teil der bei Durchführung der geplanten Studie zu erwartenden Gesamtkosten

ausmachen. Die reinen Erhebungskosten wiederum verhalten sich näherungsweise proportional zur zu erreichenden Fallzahl in der Nettostichprobe. Diese kann, wie die Überlegungen in Abschnitt 5.4 gezeigt haben, kaum niedriger angesetzt werden, wenn belastbare, ausreichend differenzierte Aussagen zu relevanten Teilgruppen von selbstständig Erwerbstätigen erreicht werden sollen. Auch können an dem in Abschnitt 4 umrissenen Programm der Kernerhebung kaum Abstriche gemacht werden, und die für die vorgenommenen Berechnungen unterstellte durchschnittliche Befragungszeit liegt eher am unteren Rand des zu dessen adäquater Umsetzung erforderlichen Zeitrahmens. Somit markieren die Aufwandsschätzungen eher die Untergrenze der Kosten, die für eine den Erkenntnisinteressen des BMAS dienliche, adäquate Umsetzung der geplanten Studie erforderlich sein dürften.

5.7 Zeitbedarf

Für die Arbeiten bis zum Beginn der Feldphase sind etwa vier Monate anzusetzen. Das umfangreiche Instrument für die Kernerhebung, aber auch das knappe Fragenprogramm für das Screening-Interview, müssen sehr sorgfältig ausgearbeitet, abgestimmt und getestet werden. Wird zur Durchführung der Untersuchung auch Datenmaterial aus einer Unternehmensdatenbank verwendet, nimmt deren Erschließung und geeignete Aufbereitung einige Zeit in Anspruch, da mit der Ziehung von Stichproben von Selbstständigen auf dieser Basis bzw. der notwendigen Umbasierung von der Unternehmens- zur Personenebene relativ wenig Erfahrungen vorhanden sind. Auch sollte in der Anlaufphase des Projekts ein Puffer für eventuell auftretende unerwartete Herausforderungen eingeplant werden.

Die zu erwartende Dauer der Feldphase ergibt sich im Wesentlichen aus den Anforderungen an das der eigentlichen Erhebung vorzuschaltenden Screening. Bei einem Design mit dualer Stichprobe ist dieses Screening für die Teilgruppe der Zielpersonen, die ausgehend von einer Unternehmensdatenbank angesprochen werden, relativ schnell zu leisten, zumal es teilweise im CAWI-Modus realisiert werden kann. Dagegen ist der Zeitbedarf für das Screening einer zufälligen Bevölkerungsstichprobe im Rahmen einer Omnibus-Befragung erheblich höher. Geht man von wochentäglich 1.000 realisierten Interviews³³ sowie Anteilen von 45 Prozent Erwerbstätigen an der erwachsenen Bevölkerung und 12 Prozent selbstständigen Zielpersonen unter den Erwerbstätigen aus, können bei einer Quote der Teilnahmebereitschaft von rund 50 Prozent etwa 140 Zielpersonen pro Woche gewonnen werden. Um allein ausgehend von einer Omnibus-Befragung eine Nettostichprobe mit der in Abschnitt 5.4 begründeten Mindestanzahl von 5.000 Zielpersonen zu gewinnen, muss also mit einer Feldzeit zur Rekrutierung von gut acht Monaten gerechnet werden. Bei einem Design mit dualer Stichprobe entfällt nur ein Anteil der zu realisierenden Gesamtzahl an Interviews der Kernerhebung auf diesen Rekrutierungsweg. Unterstellt man eine Zielgröße von 3.000, ist für das Screening per Omnibus-Befragung mit einer Feldzeit von rund fünf Monaten zu rechnen.

Es erscheint allerdings aus inhaltlichen Gründen sinnvoll, die Feldzeit für das Screening gezielt auf zwölf Monate zu verlängern, um Saisonalitäten, also systematische Schwankungen der Inzidenz vor allem kurzfristiger und sporadischer selbstständiger Beschäftigung, in einem vollständigen Jahresdurchlauf zu erfassen. Dies lässt sich realisieren, indem die Anzahl der in jedem Durchlauf der Omnibus-Befragung realisierten Screening-Interviews reduziert wird, die Screening-Fragen also nur an einen zufälligen Teil der Erwerbstätigen in der täglich realisierten Stichprobe gestellt werden, oder indem auf die Screening-Fragen in manchen Durchläufen ganz verzichtet wird. Plant man mit einer Feldzeit von einem Jahr, hat dies einen weiteren Vorteil: Man erhält eine Sicherheitsmarge für den Fall, dass der Anteil der Zielpersonen an den Erwerbstätigen oder die Rate der zur Teilnahme an der

³³ Es existieren Omnibus-Befragungen im CATI-Modus, die diese Fallzahl erreichen.

Kernbefragung bereiten Zielpersonen geringer ausfallen sollte als für die Überschlagsrechnung oben unterstellt.

Die Kernbefragung von teilnahmebereiten Zielpersonen sollte zeitnah im Anschluss an das jeweilige Screening-Interview durchgeführt werden. Dabei ist davon auszugehen, dass das umfangreiche Interview der Kernerhebung mit häufig stark eingespannten Selbstständigen eher schwer zu terminieren sein dürfte. Dennoch ist zu erwarten, dass die Kernerhebung etwa sechs Wochen nach Führung des letzten Screening-Interviews abgeschlossen sein wird. Nach Abschluss der Feldphase sind noch einmal sechs Wochen für eine abschließende Aufbereitung und Qualitätsprüfung der Daten, gegebenenfalls vorzunehmende Imputationen und die Bestimmung von Gewichten bzw. Hochrechnungsfaktoren einzuplanen. Dieser Zeitraum kann relativ kurz bemessen werden, weil die entsprechenden Arbeiten schon während der Feldzeit anhand bereits eingelaufener Daten vorbereitet werden können. Nach Abschluss der Feldphase müssen die so entwickelten und getesteten Routinen noch einmal final durchlaufen, abschließende Konsistenz- und Plausibilitätsprüfungen vorgenommen, sowie Methodenbericht und Datendokumentation erarbeitet werden. Vollständige, qualitätsgeprüfte Daten und Gewichtungsfaktoren lägen somit 19 Monate nach Projektbeginn vor.

Schließlich ist ein analytischer Endbericht zu erstellen, der die wirtschaftlichen und sozialen Lagen von Selbstständigen und insbesondere Risikofaktoren, die prekäre Verhältnisse bei dieser Beschäftigtengruppe hervorrufen, statistisch (auch mit multivariaten Verfahren) untersucht. Die erzielten Befunde sind in den Zusammenhang bereits vorliegender wissenschaftlicher Erkenntnisse zu stellen und die sich daraus ergebenden sozialpolitischen Handlungsbedarfe zu erörtern. Für die Ausarbeitung und Finalisierung des analytischen Endberichts sind drei Monate ab Vorliegen der finalen Daten zu veranschlagen. Auch hier können Analysen und statistische Routinen bereits anhand früher vorliegender, vorläufiger Daten vorbereitet werden, was die abschließende Berichtslegung beschleunigen wird.

Insgesamt ergibt sich somit ein Zeitbedarf für die Durchführung der Studie von 22 Monaten ab Auftragsvergabe.

5.8 Entwicklungsperspektiven

Die Analysemöglichkeiten auf Grundlage der Daten aus einer einmaligen Querschnittserhebung von Selbstständigen sind begrenzt. Insbesondere lassen sich Faktoren, die auf der individuellen Ebene in der erwerbsbiografischen Dynamik zu einer bestimmten, in den Daten beobachteten sozialen Lage beitragen, empirisch nicht hinreichend von den Einflüssen allgemein auf die Stichprobe einwirkender Umweltfaktoren, insbesondere der makroökonomischen Lage, trennen. Zudem lassen sich die Einflüsse relevanter, unbeobachteter individueller Merkmale, die zur Entscheidung für oder gegen die Ausübung selbstständiger Tätigkeiten beitragen und zugleich auf die Zielgröße der sozialen Lage einwirken, nicht kontrollieren, wenn nur eine einzige Beobachtung pro Person vorhanden ist. Daher wäre perspektivisch eine Überführung der geplanten Erhebung in eine Panel-Studie wünschenswert, bei der dieselben Untersuchungseinheiten mit einem gewissen zeitlichen Abstand (im Idealfall mehrfach) erneut befragt werden. Hierdurch könnten Veränderungen bei exogenen Faktoren und Zielgrößen auf der individuellen Ebene über die Zeit beobachtet werden. Die Möglichkeiten, Ursache-Wirkungs-Beziehungen statistisch zu analysieren, würden dadurch wesentlich erweitert.

Selbstverständlich ist der Ressourcenbedarf für eine Panel-Studie wegen des Aufwands für die erneute Kontaktierung (samt Nachrecherche von Kontaktdaten und Maßnahmen zur Panel-Pflege) und für die erneute Befragung um ein Vielfaches höher. Die erste Welle der Erhebung muss hinreichend groß dimensioniert sein, damit nach den zu erwartenden Panel-Ausfällen und möglichen Statuswechseln (Übergänge in Beschäftigung, Erwerbslosigkeit oder Rückzug vom Arbeitsmarkt)

noch genügend selbstständig erwerbstätige Personen wieder befragt werden können, um bei Auswertungen in der Längsschnitt-Dimension der Daten belastbare Resultate zu erzielen. Diese Bedingung wäre bei einer Ausgangsstichprobe mit netto 5.000 Fällen – und einer unterstellten Ausfallrate von etwa 50 Prozent – in einer zweiten Welle der Erhebung im Abstand von etwa einem Jahr gerade noch gegeben. Allerdings stoßen nach Teilgruppen differenzierte Auswertungen bei dieser reduzierten Basis an Fallzahlengrenzen. Ist eine Fortführung der Studie mit einer dritten Welle beabsichtigt, sollte im Rahmen der zweiten Welle eine Auffrischung der Stichprobe durch die Aufnahme von neuen Personen vorgenommen werden, die aus den dann aktuellen Grundgesamtheiten nach denselben Prinzipien ausgewählt werden wie die Teilnehmerinnen und Teilnehmer der ersten Welle.

Wird im Rahmen der geplanten Querschnitterhebung zur sozialen Lage von Selbstständigen eine hinreichend groß dimensionierte Nettostichprobe wie oben vorgeschlagen realisiert, kann die Entscheidung über die Überführung in eine Panel-Studie auch noch zu einem späteren Zeitpunkt getroffen werden. Für eine spätere Festlegung spricht, dass so auf bei der Umsetzung der ersten Erhebung gemachte Erfahrungen und angetroffene Probleme reagiert werden kann. Um die Option auf eine Folgewelle offen zu halten, sollte die Panel-Bereitschaft der Befragten jedoch bereits in der zunächst als Querschnitterhebung angelegten Studie vorsorglich abgefragt und die – von den Teilnehmerinnen und Teilnehmern jederzeit zu widerrufende – Erlaubnis zu einer eventuellen späteren erneuten Kontaktierung für die Zwecke der Studie eingeholt werden. Dabei muss aus datenschutzrechtlichen Gründen darüber informiert werden, dass im Fall der Zustimmung personenbezogene Informationen für den Zweck der erneuten Kontaktierung und die Verknüpfung der Ergebnisse der verschiedenen Befragungswellen gespeichert werden. Bei einer motivierenden Gestaltung des Interviews in der Querschnitterhebung ist mit hinreichend guten Zustimmungsqoten zu rechnen.

Eine spätere Folgewelle könnte wie die Eingangserhebung erneut als eigenständige Befragung konzipiert und durchgeführt werden. Als Alternative kommt aber auch in Frage, die realisierte Stichprobe (oder Teile davon) als Ergänzungsstichprobe in ein laufendes allgemeines Panel zu integrieren. Hierfür bietet sich insbesondere das SOEP an. Die Vorbereitungen für eine solche Übertragung müssten frühzeitig begonnen werden, da die Klärung der notwendigen Adjustierungen im Erhebungsdesign und nicht zuletzt der weiteren Finanzierung einige Zeit in Anspruch nimmt. Der Vorteil der Überführung der Studie in ein reguläres Panel wäre, dass eine statistische Datenbasis zur Bewertung der sozialen Lage selbstständig Tätiger nachhaltig entwickelt werden könnte. Dafür wäre im Vergleich zu einer vom Auftraggeber eigenständig fortgeführten Studie ein Verlust an freien Gestaltungsmöglichkeiten hinzunehmen, weil den Anforderungen der aufnehmenden Panel-Studie Rechnung getragen werden muss.

Ein anderer Ansatzpunkt, um die Auswertungsmöglichkeiten auf Grundlage der gesammelten Querschnittsdaten substanziell zu verbessern, wäre deren Verknüpfung mit den von der BA-Statistik registrierten Integrierten Erwerbsbiografien (IEB) auf der Personenebene. Auf diesem Weg ließen sich frühere Episoden abhängiger Beschäftigungsverhältnisse und gegebenenfalls des Hilfebezugs in den Bereichen Sozialgesetzbuch II (SGB II) und Sozialgesetzbuch III (SGB III) im Detail erfassen, perspektivisch aber auch die künftige erwerbsbiografische Entwicklung der Befragten abbilden, soweit sie in den von der BA registrierten Daten Spuren hinterlässt. Mit diesen Daten lassen sich zum einen Risikofaktoren in der Erwerbsbiografie, die eine aktuelle prekäre Selbstständigkeit begünstigen, genauer identifizieren. Zum anderen lassen sich mögliche erwerbsbiografische Folgen prekärer Selbstständigkeit aufzeigen und bewerten, inwieweit es Betroffenen gelingt, den Zustand durch Aufnahme einer abhängigen Beschäftigung zu verlassen. Darüber entsteht ein Anknüpfungspunkt, um die freiwillige Inanspruchnahme der Arbeitslosenversicherung durch Selbstständige zu

untersuchen, weil sich die IEB mit Daten von Selbstständigen zum Versicherungspflichtverhältnis auf Antrag aus den operativen Daten der BA verknüpfen lassen (vgl. Abschnitt 2.5).

Um die Option für die Zusammenführung der in der Erhebung gemachten Angaben mit den bei der BA zur Person des Befragten vorliegenden Sozialdaten offen zu halten, sollte die Zustimmung zu einer Verknüpfung vorsorglich eingeholt werden. Bei verknüpfungsbereiten Personen müssten zudem Angaben erfasst werden, die eine Verknüpfung mit den Sozialdaten ermöglichen. Eine eindeutige Verknüpfung wäre anhand der Sozialversicherungsnummer möglich. Allerdings ist davon auszugehen, dass viele Personen – und insbesondere Selbstständige – diese nicht ohne Weiteres angeben können. Eine gebräuchliche Alternative ist die Erfassung von Angaben wie Name, Geburtsdatum und Wohnort, die in der Kombination mit sehr hoher Wahrscheinlichkeit eine eindeutige Identifikation der befragten Zielperson in den Sozialdaten ermöglichen. Ob eine ausreichende Zustimmungsrate zustande kommt und ob gegebenenfalls die zur Verknüpfung zusätzlich erforderlichen personenbezogenen Angaben gemacht werden, lässt sich a priori schwer einschätzen. Die Freigabe des Zugangs zu personenbezogenen amtlichen Daten ist ein sensibler Punkt, und die Zustimmung dazu dürfte im CATI-Modus, weil Aufklärungsmöglichkeiten und die Vertrauensbasis in diesem Erhebungsformat beschränkter sind, spürbar geringer sein als etwa bei Erhebungen im CAPI-Modus. Auch fehlen Erfahrungen, wie es um die Verknüpfungsbereitschaft innerhalb der Zielgruppe der selbstständig Erwerbstätigen bestellt ist.

Ergibt sich dennoch eine hinreichende Verknüpfungsbereitschaft und sind zudem ausreichend Fälle mit den für eine Verknüpfung erforderlichen personenbezogenen Angaben zu erhalten, müsste die Verknüpfung im Anschluss an die Erhebung über einen Antrag auf Übermittlung von Sozialdaten für die Forschung gemäß §75 Sozialgesetzbuch X (SGB X) in die Wege geleitet werden. Weil erfahrungsgemäß die Beantragung und Realisierung der entsprechenden Verknüpfung erhebliche Zeit in Anspruch nimmt, müssen darauf basierende Untersuchungen vom Kernvorhaben der Erhebung eines Lagebilds zur sozialen Lage von Selbstständigen abgetrennt werden, um dieses in angemessener Zeit (vgl. Abschnitt 5.7) abschließen zu können.

Weitere Entwicklungsperspektiven außerhalb des Kernvorhabens ergeben sich daraus, dass im Anschluss spezielle Teilgruppen in der Population der selbstständig Tätigen, für die im Rahmen einer bevölkerungsrepräsentativ angelegten Studie fallzahlenbedingt keine belastbaren Aussagen erzielt werden konnten, in eigenständigen Studien näher untersucht werden könnten. Dafür müssten dann auch selektive Zugangswege, wie sie in Abschnitt 3.2.3 skizziert wurden, beschränkt werden. Die geplante Querschnittserhebung wäre in diesem Kontext in zweierlei Hinsicht nützlich. Erstens steht damit ein erprobtes Erhebungsinstrument zur Erfassung der sozialen Lage von Selbstständigen zur Verfügung. Die Weiterwendung des Instruments ermöglicht ein Benchmarking der Befunde einer anschließenden Studie für eine spezielle Teilpopulation. Zweitens lassen sich sozio-demografische Merkmale dieser Teilpopulation, die im Rahmen der bevölkerungsrepräsentativen Befragung gewonnen wurden, dafür nutzen, bei einer weiter gefassten Erhebung mit nicht zufälliger Stichprobenbasis auftretende Selektivitäten zu erkennen.

Literaturverzeichnis

- Baresel, K., H. Eulitz, U. Fachinger, M. M. Grabka, C. Halbmeier, H. Künemund, A. Lozano Ancántara und C. Vogel (2021). Hälfte aller Erbschaften und Schenkungen geht an die reichsten zehn Prozent aller Begünstigten. DIW Wochenbericht 5/2021, 63-71. Deutsches Institut für Wirtschaftsforschung Berlin. https://www.diw.de/documents/publikationen/73/diw_01.c.809828.de/21-5-1.pdf (zuletzt aufgerufen am 12. August 2022).
- Bertschek, I. und D. Erdsiek (2020). Soloselbstständigkeit in der Corona-Krise, Digitalisierung hilft bei der Bewältigung der Krise. ZEW-Kurzexpertise Nr. 20-08. Leibniz-Zentrum für Europäische Wirtschaftsforschung Mannheim. https://ftp.zew.de/pub/zew-docs/ZEWKurzexpertisen/ZEW_Kurzexpertise2008.pdf (zuletzt aufgerufen am 12. August 2022).
- Beznoska, M. (2017). Vermögen der Selbstständigen: Mikrodatenanalyse mit dem Sozio-ökonomischen Panel für das Jahr 2012. IW-Report Nr. 20. Institut der deutschen Wirtschaft Köln. <https://www.iwkoeln.de/studien/martin-beznoska-mikrodatenanalyse-mit-dem-sozio-oekonomischen-panel-fuer-das-jahr-2012.html> (zuletzt aufgerufen am 12. August 2022).
- Bonin, H., A. Krause-Pilatus und U. Rinne (2022). Selbstständige Erwerbstätigkeit in Deutschland (Aktualisierung 2022). BMAS Forschungsbericht 601. Bundesministerium für Arbeit und Soziales Berlin. <https://www.bmas.de/DE/Service/Publikationen/Forschungsberichte/fb-601-selbststaendige-erwerbstaetigkeit-in-deutschland.html> (zuletzt aufgerufen am 12. August 2022).
- Bonin, H. und U. Rinne (2017). Omnibusbefragung zur Verbesserung der Datenlage neuer Beschäftigungsformen, Kurzexpertise im Auftrag des Bundesministeriums für Arbeit und Soziales. IZA Research Report No. 80. Forschungsinstitut zur Zukunft der Arbeit Bonn. https://docs.iza.org/report_pdfs/iza_report_80.pdf (zuletzt aufgerufen am 12. August 2022).
- Börsch-Supan, A., M. Coppola, L. Essig, A. Eymann und D. Schunk (2009). The German SAVE study Design and Results 2009. MEA Studie Nr. 6. Mannheim Research Institute for the Economics of Aging, Universität Mannheim.
- Bührmann, A. D., U. Fachinger und E. M. Welskoop-Deffaa (Hrsg.) (2018). Hybride Erwerbsformen: Digitalisierung, Diversität und sozialpolitische Gestaltungsoptionen. Wiesbaden.
- Bundesregierung (2021). Lebenslagen in Deutschland. Der Sechste Armuts- und Reichtumsbericht der Bundesregierung. Berlin. <https://www.armuts-und-reichtumsbericht.de/SharedDocs/Downloads/Berichte/sexster-armuts-reichtumsbericht.pdf?blob=publicationFile&v=6> (zuletzt aufgerufen am 12. August 2022).
- Fowler, F. J. (2014). Survey Research Methods. Los Angeles.
- Gensicke, M. und N. Tschersich (2014). Methodenexperiment im Rahmen der BIBB/BAuA-Erwerbstätigenbefragung 2011/2012: Ein Vergleich von CATI, CAPI und CAWI. Abschlussbericht zum Projekt F 2296 im Auftrag der BAuA. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund, Berlin und Dresden. <https://www.baua.de/DE/Angebote/Publikationen/Berichte/F2296.pdf?blob=publicationFile&v=2> (zuletzt aufgerufen am 12. August 2022)

- Grabka, M. M. und C. Westermeier (2015). Editing and Multiple Imputation of Item Non-response in the Wealth Module of the German Socio-Economic Panel. SOEP Survey Papers 272, Series C – Data Documentations, Deutsches Institut für Wirtschaftsforschung Berlin.
https://www.diw.de/documents/publikationen/73/diw_01.c.570790.de/diw_ssp0272.pdf
(zuletzt aufgerufen am 12. August 2022)
- Gruber, J. (2019). Hybride Erwerbsformen – Geschlechter- und branchenspezifische Unterschiede. Wirtschaftsdienst 99(7), 515-518. <https://www.wirtschaftsdienst.eu/inhalt/jahr/2019/heft/7/beitrag/hybride-erwerbsformen-geschlechter-und-branchenspezifische-unterschiede.html>
(zuletzt aufgerufen am 12. August 2022).
- Heien, T. und M. Krämer (2021). Alterssicherung in Deutschland 2019 (ASID 2019) – Zusammenfassender Bericht. BMAS Forschungsbericht 572/Z. Bundesministerium für Arbeit und Soziales Berlin. https://www.bmas.de/SharedDocs/Downloads/DE/Publikationen/Forschungsberichte/fb-572-alterssicherung-in-deutschland-2019.pdf?__blob=publicationFile&v=2 (zuletzt aufgerufen am 12. August 2022).
- Jäckle, A., P. Lynn und J. Burton (2015). Going Online with a Face-to-Face Household Panel: Effects of a Mixed Mode Design on Item and Unit Non-Response. Survey Research Methods 9(1), 57-70.
- Jahn, E. und A. Springer (2013). Arbeitslosenversicherung: Auch Selbstständige nehmen Unterstützung in Anspruch. IAB-Kurzbericht 26/2013. Institut für Arbeitsmarkt- und Berufsforschung Nürnberg. <https://doku.iab.de/kurzber/2013/kb2613.pdf> (zuletzt aufgerufen am 12. August 2022).
- Kritikos, A. S., D. Graeber und J. Seebauer (2020). Corona-Pandemie wird zur Krise für Selbstständige. DIW aktuell Nr. 47, 12. Juni 2020. Deutsches Institut für Wirtschaftsforschung Berlin. https://www.diw.de/documents/publikationen/73/diw_01.c.791679.de/diw_aktuell_47.pdf
(zuletzt aufgerufen am 12. August 2022).
- Lüdtke, D. und J. Schupp (2017). Wechsel von persönlichen Interviews zu webbasierten Interviews in einem laufenden Haushaltspanel. In: S. Eifler und F. Faulbaum (Hrsg.), Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung. Schriftenreihe der ASI - Arbeitsgemeinschaft Sozialwissenschaftlicher Institute. Wiesbaden.
https://doi.org/10.1007/978-3-658-15834-7_6 (zuletzt aufgerufen am 12. August 2022).
- Lingo, E. L. und S. J. Tepper (2013), Looking Back, Looking Forward: Arts-Based Careers and Creative Work, Work and Occupations, 40(4), 337-363.
- Menold, N. (2015). Befragung. GESIS Survey Guidelines – Schriftlich-postalische Befragung, Version 1.0. https://www.gesis.org/fileadmin/upload/SDMwiki/Archiv/Schriftlich-postalische_Befragung_Menold_12092014_1.0.pdf (zuletzt aufgerufen am 12. August 2022).
- Metzger, G. (2021). Enger Zusammenhang zwischen Risikobereitschaft und Selbstständigkeit, Fokus Volkswirtschaft Nr. 317, KfW Research Frankfurt. <https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Fokus-Volkswirtschaft/Fokus-2021/Fokus-Nr.-317-Februar-2021-Risikobereitschaft.pdf> (zuletzt aufgerufen am 12. August 2022).

- Metzger, G. (2020). Blitzbefragung: Auswirkungen der Corona-Pandemie trifft Selbstständige hart. Fokus Volkswirtschaft Nr. 282, KfW Research Frankfurt. <https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Fokus-Volkswirtschaft/Fokus-2020/Fokus-Nr.-282-April-2020-Blitzumfrage-Gruenderplattform.pdf> (zuletzt aufgerufen am 12. August 2022).
- Metzger, G. (2021). „Junge“ Selbstständige: branchenbedingt trifft Corona-Krise Frauen härter als Männer. Fokus Volkswirtschaft Nr. 324, KfW Research Frankfurt. <https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Fokus-Volkswirtschaft/Fokus-2021/Fokus-Nr.-324-Maerz-2021-Junge-Selbststaendige.pdf> (zuletzt aufgerufen am 12. August 2022).
- Niehues, J. und J. Pimpertz (2012). Alterssicherung der Selbstständigen in Deutschland. IW-Trends 3/2012. Institut der deutschen Wirtschaft Köln. <https://www.iwkoeln.de/studien/judith-niehues-jochen-pimpertz-alterssicherung-der-selbststaendigen-in-deutschland-91223.html> (zuletzt aufgerufen am 12. August 2022).
- Oberfichtner, M. (2019). Arbeitslosenversicherung für Existenzgründer: Unterschiedliche Leistungen trotz gleicher Beiträge. IAB-Kurzbericht 01/2019. Institut für Arbeitsmarkt- und Berufsforschung Nürnberg. <https://doku.iab.de/kurzber/2019/kb0119.pdf> (zuletzt aufgerufen am 12. August 2022).
- Pahnke, A., S. Schneck und H.-J. Wolter (2019). Persistenz von Selbstständigen in der Grundsicherung. IfM-Materialien Nr. 273. Institut für Mittelstandsforschung Bonn. https://www.ifm-bonn.org/fileadmin/data/redaktion/publikationen/ifm_materialien/dokumente/IfM-Materialien-273-2019.pdf (zuletzt aufgerufen am 12. August 2022).
- Promberger, M., K. Jahn, B. Schels, J. Allmendinger und S. Stuth (2018). Existiert ein verfestigtes „Prekariat“? Prekäre Beschäftigung, ihre Gestalt und Bedeutung im Lebenslauf und die Konsequenzen für die Strukturierung sozialer Ungleichheit. Working Paper Forschungsförderung, Nummer 085. Hans-Böckler-Stiftung Düsseldorf. https://www.boeckler.de/pdf/p_fofoe_WP_085_2018.pdf (zuletzt aufgerufen am 12. August 2022).
- Rodgers, G. und J. Rodgers (1989). Precarious Jobs in Labour Market Regulation: The Growth of Atypical Employment in Western Europe. International Labour Organization, Genf.
- Rosemann, M. und A. Koch (2012). Zur Weiterentwicklung der sozialen Sicherungssysteme (Arbeitslosenversicherung) für Soloselbstständige in der Kreativwirtschaft. In: „Soziale Sicherung für Soloselbstständige in der Kreativwirtschaft“. WISO Diskurs März 2012, 5-26. Friedrich-Ebert-Stiftung Bonn. <https://library.fes.de/pdf-files/wiso/08940.pdf> (zuletzt aufgerufen am 12. August 2022).
- Schnell, R. (2019). Survey-Interviews. Methoden standardisierter Befragung. Studienskripte zur Soziologie. 2. Auflage Wiesbaden.
- Schröder, J. (2015). GESIS Survey Guidelines – Persönlich-mündliche Befragung. Version 1.1. https://www.gesis.org/fileadmin/upload/SDMwiki/Pers.-muendl.Befragung_Schroeder_08102015_1.1.pdf (zuletzt aufgerufen am 12. August 2022)

- Schröder, C., C. Bartels, K. Göbler, M. M. Grabka und J. König (2020). MillionärInnen unter dem Mikroskop: Datenlücke bei sehr hohen Vermögen geschlossen – Konzentration höher als bisher ausgewiesen. DIW Wochenbericht 29/2020, 511-521. Deutsches Institut für Wirtschaftsforschung Berlin. https://www.diw.de/documents/publikationen/73/diw_01.c.793785.de/20-29-1.pdf (zuletzt aufgerufen am 12. August 2022).
- Seebauer, J., A. Kritikos und D. Graeber (2021). Warum vor allem weibliche Selbstständige Verliererinnen der Covid-19-Krise sind. DIW Wochenbericht 15/2021, 262-269. https://www.diw.de/documents/publikationen/73/diw_01.c.815784.de/21-15-3.pdf (zuletzt aufgerufen am 12. August 2022).
- Schulze Buschhoff, K., W. Coenen und J. Schippers (2017), Solo-Selbstständigkeit – Eine prekäre Beschäftigungsform? WSI-Mitteilungen 1/2017, 54-61. Wirtschafts- und Sozialwissenschaftliches Institut der Hans-Böckler-Stiftung Düsseldorf. https://www.wsi.de/data/wsimit_2017_01_schulze_buschhoff.pdf (zuletzt aufgerufen am 12. August 2022).
- Schulze Buschhoff, K. und H. Emmeler (2021). Selbstständige in der Corona-Krise: Ergebnisse aus der HBS-Erwerbspersonenbefragung, Wellen 1 bis 5. WSI Policy Brief Nr. 60. Wirtschafts- und Sozialwissenschaftliches Institut der Hans-Böckler-Stiftung Düsseldorf. https://www.boeckler.de/fpdf/HBS-008112/p_wsi_pb_60_2021.pdf (zuletzt aufgerufen am 12. August 2022).
- Standing, G. (2011). The Precariat: The New Dangerous Class. London.
- Westerveld, M. (2012), The „New“ Self-employed: An Issue for Social Policy? European Journal of Social Security 14(3), 156-173.
- Vosko, L. F., M. Macdonald und I Campbell (2009). Gender and the Contours of Precarious Employment. London/New York.

Diese Publikation wird im Rahmen der Öffentlichkeitsarbeit des Bundesministeriums für Arbeit und Soziales kostenlos herausgegeben. Sie darf weder von Parteien noch von Wahlbewerbern oder Wahlhelfern während des Wahlkampfes zum Zwecke der Wahlwerbung verwendet werden. Dies gilt für Europa-, Bundestags-, Landtags- und Kommunalwahlen. Missbräuchlich ist insbesondere die Verteilung auf Wahlveranstaltungen, an Informationsständen der Parteien sowie das Einlegen, Aufdrucken oder Aufkleben parteipolitischer Informationen oder Werbemittel. Untersagt ist gleichfalls die Weitergabe an Dritte zum Zwecke der Wahlwerbung. Unabhängig davon, wann, auf welchem Weg und in welcher Anzahl diese Publikation dem Empfänger zugegangen ist, darf sie auch ohne zeitlichen Bezug zu einer bevorstehenden Wahl nicht in einer Weise verwendet werden, die als Parteinahme der Bundesregierung zugunsten einzelner politischer Gruppen verstanden werden könnte. Außerdem ist diese kostenlose Publikation - gleichgültig wann, auf welchem Weg und in welcher Anzahl diese Publikation dem Empfänger zugegangen ist - nicht zum Weiterverkauf bestimmt.

Alle Rechte einschließlich der fotomechanischen Wiedergabe und des auszugsweisen Nachdrucks vorbehalten.