

9 Methoden zur Erhebung von Daten über Innovationen im Unternehmenssektor

Auf der Grundlage der in den vorangegangenen Kapiteln vorgestellten Konzepte und Definitionen enthält dieses Kapitel Leitlinien zur Methodik für die Erhebung von Daten über Innovationen im Unternehmenssektor. Die Leitlinien richten sich an Produzenten von Innovationsstatistiken sowie an fortgeschrittene Nutzer, die den Prozess der Entstehung von Innovationsdaten verstehen müssen. Im Mittelpunkt des Kapitels steht die Verwendung von Innovationserhebungen, um Daten über verschiedene Aspekte innovationsbezogener Aktivitäten und Ergebnisse innerhalb von Unternehmen sowie sonstige damit zusammenhängende Informationen zu sammeln. Zudem werden weitere potenzielle Quellen genannt. Die in diesem Kapitel formulierten Leitlinien beziehen sich auf den gesamten Zyklus einer Datenerhebung, d. h. die Festlegung der Ziele und des Erfassungsbereichs von Innovationserhebungen, die Bestimmung der Grundgesamtheit, das Fragebogendesign, die Stichprobenverfahren, die Datenerhebungsmethoden und die Erhebungskonzepte, die Aufbereitung der Erhebungsdaten und die Veröffentlichung statistischer Ergebnisse.

9.1. Einleitung

9.1. Dieses Kapitel enthält Leitlinien zur Methodik für die Erhebung von Daten über Innovationen im Unternehmenssektor. Wie in Kapitel 2 erörtert, sind methodische Leitlinien für die Erhebung von Innovationsdaten ein wesentlicher Bestandteil des Rahmens zur Messung von Innovationen. Innovationsdaten lassen sich durch objektorientierte Methoden wie z. B. neue Produktankündigungen im Internet oder in Fachzeitschriften (Kleinknecht, Reijnen und Smits, 1993) sowie aus Expertisen über Innovationen (Harris, 1988) gewinnen. Andere Quellen für Innovationsdaten sind Jahresberichte und Websites von Unternehmen, Personen- und Haushaltserhebungen über das Bildungsniveau der Arbeitskräfte, Berichte an regionale, nationale und supranationale Organisationen, die Forschung und experimentelle Entwicklung (FuE) oder Innovationen finanzieren, Berichte an Organisationen, die Innovationspreise vergeben, Wissenstransferstellen von Universitäten, die Daten über die von Unternehmen finanzierte Auftragsforschung und die Lizenzierung des geistigen Eigentums von Universitäten erfassen, sowie Unternehmensregisterangaben, administrative Quellen und Erhebungen über Unternehmensaktivitäten, FuE und den Einsatz von Informations- und Kommunikationstechnologien (IKT). Viele dieser existierenden und potenziellen künftigen Quellen weisen Merkmale von Big Data auf, d. h., sie sind zu groß oder zu komplex für herkömmliche Instrumente und Methoden der Datenverarbeitung.

9.2. Obwohl diese Datenquellen für verschiedene Zwecke nützlich sind, gibt es dennoch Grenzen ihrer Verwendbarkeit. Viele bieten weder auf Wirtschaftszweig- noch auf nationaler Ebene eine repräsentative Erfassung von Innovationen, weil die Daten auf Selbstselektion basieren: Berücksichtigt werden nur Unternehmen, die eine Produktankündigung machen, FuE-Förderung beantragen oder Wissen von Universitäten lizenzieren. Informationen aus Unternehmensregistern, Personen- und Haushaltserhebungen sowie Erhebungen über Unternehmensaktivitäten und FuE sind oft unvollständig, da sie nur einen Aspekt einer Innovation abdecken. Auch die Jahresberichte und Websites von Unternehmen erfassen die Innovationsaktivitäten nicht unbedingt auf konsistente Weise, wengleich Web-Scraping-Verfahren eine Automatisierung der Suche nach Innovationsaktivitäten in online gestellten Dokumenten ermöglichen und in Zukunft eine zunehmend wertvolle Quelle für Innovationsdaten darstellen können. Zwei weitere Einschränkungen sind, dass keine dieser Quellen konsistente und vergleichbare Daten zum gesamten Spektrum der von allen Unternehmen durchgeführten Innovationsstrategien und -aktivitäten liefert (wie in den Kapiteln 3 bis 8 erörtert), und dass sich einige der Quellen nicht miteinander verknüpfen lassen. Die einzige Möglichkeit, einen vollständigen Satz konsistenter und miteinander verknüpfbarer Daten zu erhalten, besteht derzeit darin, eine gesonderte Innovationserhebung durchzuführen, die auf einem Unternehmensregister basiert.

9.3. Ziel einer Innovationserhebung ist es, qualitativ hochwertige Daten über Innovationen in Unternehmen von maßgebenden Personen, wie den Geschäftsführer*innen oder leitenden Führungskräften zu erhalten. Ob dieses Ziel erreicht wird, hängt von einer Vielzahl von Faktoren ab, darunter dem Erfassungsgrad der Grundgesamtheit, der Häufigkeit der Datenerhebung, der Fragenformulierung und dem Fragebogendesign und den entsprechenden Pre-Tests, der Erstellung des Stichprobenrahmens, den zur Durchführung der Erhebung verwendeten Methoden (einschließlich der Identifizierung einer geeigneten Antwortperson innerhalb der befragten Einheit) sowie der Datenaufbereitung. All diese Aspekte sind für die nationalen Statistikämter wie auch für internationale Organisationen und Forschende relevant, die Daten zu Innovationsaktivitäten anhand von Erhebungen gewinnen und analysieren wollen.

9.4. Die von den nationalen Statistikämtern im Rahmen der Erstellung nationaler Unternehmensstatistiken durchgeführten Innovationserhebungen müssen den jeweiligen Praktiken für Fragebogen- und Erhebungsdesign folgen. Die in diesem Kapitel dargelegten Empfehlungen beziehen sich auf empfehlenswerte Praktiken, die für die meisten nationalen Statistikämter umsetzbar sein sollten. Erhebungen, die außerhalb des amtlichen statistischen Systems vorgenommen werden, z. B. von internationalen Organisationen oder Hochschulen, können von den in diesem Kapitel formulierten Empfehlungen profitieren (OECD, 2015a). Begrenzte Ressourcen und rechtliche Einschränkungen können es Organisationen jedoch erschweren, alle empfehlenswerten Praktiken umzusetzen.

9.5. Die Entscheidung, welche Arten von Daten in einer Erhebung gesammelt werden, sollte in Absprache mit den Datennutzern wie Expert*innen, die politische Maßnahmen entwickeln und analysieren, Manager*innen und Berater*innen von Unternehmen, Wissenschaftler*innen und sonstigen Akteuren getroffen werden. Die von den nationalen Statistikämtern durchgeführten Erhebungen werden hauptsächlich von politischen Entscheidungsträger*innen und Einrichtungen, die die Politik beraten, genutzt. Daher sollte die Auswahl der Fragen nach Absprache mit den für Innovation und Unternehmensentwicklung zuständigen Ministerien und Behörden erfolgen. Von Wissenschaftler*innen erstellte Erhebungen könnten ebenfalls von Konsultationen mit staatlichen Stellen oder Unternehmen profitieren.

9.6. Der Zweck der Datenerhebung – z. B. die Erstellung nationaler oder regionaler Indikatoren oder Verwendung in der Forschung – wirkt sich entscheidend auf die Wahl der Erhebungsmethode aus. Die Stichprobe kann kleiner ausfallen, wenn lediglich auf Indikatoren auf nationaler Ebene abgezielt wird. Eine größere Stichprobe ist hingegen erforderlich, wenn die Nutzer Daten über Teilgesamtheiten, Längsschnittdaten oder Daten über seltene Innovationsphänomene benötigen. Zudem hat der Erhebungszweck einen starken Einfluss auf die Arten der Fragen, die in den Fragebogen aufgenommen werden sollen.

9.7. Dieses Handbuch enthält mehr Vorschläge für Fragen zum Thema Innovation, als in einer einzigen Erhebung berücksichtigt werden können. In den Kapiteln 3 bis 8 und in Kapitel 10 werden wichtige Fragen für eine regelmäßige und zusätzliche Fragen für eine gelegentliche Aufnahme in die Fragebögen von Innovationserhebungen vorgeschlagen. Zusätzliche Fragen, die auf ergänzenden Empfehlungen oder anderen Abschnitten des Handbuchs basieren, können in Sondermodulen zu bestimmten Themen oder in gesonderten, speziellen Erhebungen berücksichtigt werden. Die in diesem Kapitel enthaltenen Empfehlungen sind für Vollerhebungen, spezielle Erhebungen und für Innovationsmodule im Rahmen sonstiger Erhebungen relevant.

9.8. Dieses Kapitel befasst sich eingehender mit Empfehlungen zu Erhebungsmethoden als frühere Ausgaben dieses Handbuchs. Viele Leser*innen der nationalen Statistikämter sind damit vertraut und brauchen in vielen Punkten keine genaueren Leitlinien. Diese Ausgabe soll aber neben nationalen Statistikämtern auch anderen Produzenten und Nutzern von Innovationsdaten weltweit dienen. Letztere könnten daher die in diesem Kapitel gegebenen Erläuterungen für ihre Arbeit nützlich finden. Weitere Quellen für allgemeine Leitlinien zu Erhebungen im Unternehmenssektor sind u. a. Willeboordse (Hrsg.) (1997) und Snijkers et al. (Hrsg.) (2013). Die Online-Ausgabe dieses Handbuchs wird durch einschlägige Links zu aktuellen und jüngsten Erhebungsverfahren sowie zu Beispielen für experimentelle Ansätze mit neuen Datenerhebungsmethoden ergänzt (<http://oe.cd/oslomanual>).

9.9. Das Kapitel ist wie folgt aufgebaut: In Abschnitt 9.2 werden die Grundgesamtheit und sonstige wesentliche Merkmale besprochen, die für Innovationserhebungen von Bedeutung sind. Fragebogendesign und Fragenformulierung werden in Abschnitt 9.3 behandelt. Danach werden verschiedene methodische Aspekte einer Erhebung erläutert, u. a. die Stichprobenaus-

wahl (Abschnitt 9.4), die Datenerhebungsmethoden (Abschnitt 9.5), das Erhebungskonzept (Abschnitt 9.6) und die Datenaufbereitung (Abschnitt 9.7). Das Kapitel schließt mit einer kurzen Darstellung verschiedener Aspekte der Veröffentlichung und Verbreitung von Ergebnissen aus Innovationserhebungen (Abschnitt 9.8).

9.2. Grundgesamtheit und weitere wesentliche Merkmale einer Erhebung

9.2.1. Grundgesamtheit

9.10. Abdeckungsbereich von Innovationserhebungen ist der in Kapitel 2 und OECD (2015b) definierte Unternehmenssektor. Er umfasst

- alle gebietsansässigen Unternehmen mit eigener Rechtspersönlichkeit, ungeachtet der Gebietsansässigkeit ihrer Anteilseigner. Dazu zählen Einheiten, die in der Lage sind, einen Gewinn oder sonstigen finanziellen Nutzen für ihre Eigentümer zu erwirtschaften, die gesetzlich als rechtlich selbstständig von ihren Eigentümern anerkannt sind und die für die Zwecke der Marktproduktion zu wirtschaftlich signifikanten Preisen errichtet wurden. Dazu gehören finanzielle und nichtfinanzielle Kapitalgesellschaften sowie Personengesellschaften und Einzelunternehmen.
- die rechtlich unselbstständigen Zweigniederlassungen gebietsfremder Unternehmen, die als gebietsansässig und Teil dieses Sektors gelten, weil sie auf langfristiger Basis als Produzenten in diesem Wirtschaftsgebiet tätig sind.
- alle gebietsansässigen Organisationen ohne Erwerbszweck, die Marktproduzenten von Waren oder Dienstleistungen oder für andere Unternehmen tätig sind. Dazu gehören unabhängige Forschungsinstitute, Kliniken und sonstige Einrichtungen, deren Haupttätigkeit die Produktion von Waren und Dienstleistungen ist, die zu Preisen verkauft werden, die ihre vollen wirtschaftlichen Kosten decken. Ebenfalls dazu zählen Einheiten, die von Unternehmensverbänden kontrolliert und über Beiträge und Gebühren finanziert werden.

9.11. Dem Unternehmenssektor gehören sowohl private Unternehmen (ob börsennotiert oder nicht) als auch staatlich kontrollierte Unternehmen (die als „öffentliche Unternehmen“ bezeichnet werden) an. Bei öffentlichen Unternehmen wird die Abgrenzung zwischen dem Unternehmenssektor und dem Staatssektor danach definiert, in welchem Umfang die Einheit in der Marktproduktion tätig ist. Wenn die Haupttätigkeit einer Einheit in der Produktion von Waren oder Dienstleistungen zu wirtschaftlich signifikanten Preisen besteht, wird sie als Unternehmen betrachtet.

9.12. Gemäß der Definition im System der Volkswirtschaftlichen Gesamtrechnungen der Vereinten Nationen (SNA 2008) (Europäische Kommission et al., 2009) ist eine Einheit in dem Wirtschaftsgebiet gebietsansässig, mit dem es die engste Verbindung aufweist und in dem sie mindestens ein Jahr wirtschaftliche Tätigkeiten ausübt. Ein Wirtschaftsgebiet kann ein geografisches Gebiet oder Hoheitsgebiet sein, für das Statistiken erstellt werden, wie z. B. ein Land, ein Bundesland bzw. eine Provinz oder eine Region. Der wirtschaftliche Interessenschwerpunkt von Unternehmen liegt normalerweise in dem Land, in dem sie rechtlich gegründet wurden und registriert sind. Sie können in anderen Ländern gebietsansässig sein als ihre Anteilseigner, und Tochterunternehmen können in anderen Ländern gebietsansässig sein als ihre Muttergesellschaften.

9.13. Die wichtigsten Merkmale der Grundgesamtheit, die bei der Erstellung einer Stichproben- oder Vollerhebung berücksichtigt werden müssen, sind die Art der statistischen Einheit, die wirtschaftliche Haupttätigkeit der Einheit, die Größe der Einheit und der geografische Standort der Einheit.

9.2.2. Statistische Einheiten und Berichtseinheiten

9.14. Unternehmen organisieren ihre Innovationsaktivitäten auf verschiedenen Ebenen, um ihre Ziele zu erreichen. Strategische Entscheidungen, die die Finanzierung und die Ausrichtung der Innovationsanstrengungen betreffen, werden oft auf Unternehmensebene getroffen. Sie können aber auch auf Ebene der Unternehmensgruppe gefällt werden, ungeachtet nationaler Grenzen. Es ist ebenfalls möglich, dass Manager*innen unterhalb der Ebene des Unternehmens (d. h. auf Ebene des Betriebs oder der fachlichen Einheit [FE]) Entscheidungen treffen, die für Innovationen relevant sind.

9.15. Diese Entscheidungen können auf grenzüberschreitender Ebene stattfinden, insbesondere im Fall multinationaler Unternehmen (MNU). Dies kann es erschweren, die für die Entscheidungsfindung verantwortlichen Personen zu identifizieren und zu befragen, vor allem wenn die nationalen Statistikämter oder sonstige Datenerfassungsstellen nur die Befugnis haben, Informationen von inländischen Einheiten zu sammeln.

Statistische Einheit

9.16. Eine **statistische Einheit** ist eine Einheit, *über die* Informationen eingeholt werden und die Gegenstand von Statistiken ist; d. h., sie ist die institutionelle Einheit, die für den vorgesehenen Zweck der Erhebung von Innovationsstatistiken von Interesse ist. Eine statistische Einheit kann eine **Beobachtungseinheit** sein, über die Informationen eingeholt und Statistiken erstellt werden, oder eine **Analyseeinheit**, die durch Aufteilung oder Kombination von Beobachtungseinheiten mithilfe von Schätzungen oder Imputationen geschaffen wird, um detailliertere oder homogenere Daten zu liefern als andernfalls möglich wäre (VN, 2007; OECD, 2015b).

9.17. Die Notwendigkeit zur Abgrenzung statistischer Einheiten entsteht bei großen und komplexen wirtschaftlichen Einheiten, deren Tätigkeiten in verschiedene Wirtschaftszweige fallen oder deren Einheiten in verschiedenen geografischen Gebieten liegen. In Abhängigkeit von ihren Eigentums- und Kontrollverhältnissen, der Homogenität ihrer wirtschaftlichen Tätigkeit sowie ihrem Standort werden verschiedene Arten von statistischen Einheiten unterschieden – Unternehmensgruppen, Unternehmen, Betriebe (eine Einheit, die sich an einem einzigen Standort befindet und in der eine einzige Produktionstätigkeit ausgeübt wird) und fachliche Einheiten (Teil einer Einheit, in der nur eine einzige Art von Produktionstätigkeit stattfindet) (vgl. OECD [2015b: Kasten 3.1] für Einzelheiten). Die Wahl der statistischen Einheit und der zur Datenerhebung verwendeten Methodik wird stark durch den Zweck der Innovationstatistiken, die Existenz von Aufzeichnungen über die Innovationstätigkeit innerhalb der Einheit und die Fähigkeit der Antwortpersonen, die gewünschten Informationen zu liefern, beeinflusst.

9.18. Als statistische Einheit in Erhebungen im Unternehmenssektor dient in der Regel das Unternehmen, das gemäß der Definition im SNA die kleinste Kombination rechtlicher Einheiten ist, die über Autonomie bei Finanz- und Anlageentscheidungen sowie Befugnis und Verantwortung für die Zuweisung von Ressourcen zur Produktion von Waren und Dienstleistungen verfügt (Europäische Kommission et al., 2009; OECD, 2015b: Kasten 3.1).

9.19. Deskriptive Identifizierungsvariablen sollten für alle statistischen Einheiten in der Grundgesamtheit einer Innovationserhebung vorliegen. Diese Variablen sind in der Regel in statistischen Unternehmensregistern enthalten und umfassen für jede statistische Einheit einen Identifizierungscode, den Standort, die Art der durchgeführten wirtschaftlichen Tätigkeit und die Größe der Einheit. Zusätzliche Informationen zur wirtschaftlichen oder rechtlichen Organisation einer statistischen Einheit sowie zu ihren Eigentumsverhältnissen und ihrer Sektorzugehörigkeit (öffentlicher oder privater Status) können hilfreich sein, um den Erhebungsprozess effektiver und effizienter zu gestalten.

Berichtseinheiten

9.20. Die **Berichtseinheit** (d. h. die „Ebene“ innerhalb des Unternehmens, von der die erforderlichen Daten eingeholt werden) ist aufgrund der jeweiligen institutionellen Strukturen, rechtlichen Rahmenbedingungen für die Datenerhebung, Traditionen, nationalen Prioritäten, vorhandenen Ressourcen für die Erhebung und individuellen Übereinkünften mit den befragten Unternehmen von Land zu Land (und möglicherweise innerhalb eines Landes) unterschiedlich ausgefallen. Die Berichtseinheit muss daher nicht deckungsgleich mit der gewünschten statistischen Einheit sein. Es kann erforderlich sein, die von den Berichtseinheiten gelieferten Informationen (durch Interpolation oder Schätzung) zu kombinieren, aufzuteilen oder zu ergänzen, um sie an die gewünschte statistische Einheit anzupassen.

9.21. Unternehmen können aus mehreren Betrieben oder fachlichen Einheiten bestehen, bei vielen kleinen und mittleren Unternehmen (KMU) sind Betrieb und Unternehmen aber meist identisch. Wenn ein Unternehmen verschiedene wirtschaftliche Tätigkeiten ausübt, kann es für regionalpolitische Interessen notwendig sein, Daten auf Ebene der fachlichen Einheit oder des Betriebs zu erheben. Bei der Stichprobenauswahl von Betrieben oder fachlichen Einheiten ist jedoch darauf zu achten, dass sie bei der Datenaggregation nicht mehrfach erfasst werden.

9.22. Wenn Informationen nur auf höheren Aggregationsniveaus – wie der Unternehmensgruppe – verfügbar sind, kann es notwendig sein, dass nationale Statistikämter mit diesen Einheiten zusammenarbeiten müssen, um disaggregierte Daten zu erhalten, indem sie beispielsweise eine Aufschlüsselung der Daten nach Hoheitsgebiet oder Wirtschaftstätigkeit beantragen. Dies ermöglicht eine bessere Harmonisierung mit anderen Wirtschaftsstatistiken.

9.23. Die Unternehmensgruppe kann als Berichtseinheit von herausragender Bedeutung sein, wenn von einer zentralen Verwaltungsstelle Fragebögen ausgefüllt oder Antworten genehmigt werden. Im Fall von Holding-Gesellschaften kann eine Reihe verschiedener Ansätze verfolgt werden. Beispielsweise kann die Holding-Gesellschaft gebeten werden, über die Innovationsaktivitäten der Unternehmen in bestimmten Wirtschaftszweigen zu berichten. Ein weiterer Ansatz besteht darin, den Fragebogen oder relevante Abschnitte an andere Teile des Unternehmens weiterzuleiten.

9.24. Auch wenn es für politische Entscheidungen oder aus praktischen Gründen Innovationsdaten auf Ebene der Betriebe, fachlichen Einheiten und Unternehmensgruppen bedarf, wird empfohlen, die Daten nach Möglichkeit auf Unternehmensebene zu erheben, um internationale Vergleiche anstellen zu können. Wenn dies nicht möglich ist, sollte besondere Aufmerksamkeit auf die Erhebung von Daten über Innovationsaktivitäten und -aufwendungen sowie von Informationen über Innovationsverflechtungen gelegt werden, da diese Elemente auf den verschiedenen Aggregationsebenen u. U. nicht additiv sind, insbesondere im Fall von multinationalen Unternehmen. Darüber hinaus können Innovationsaktivitäten Teil komplexer globaler Wertschöpfungsketten sein, die weit verteilte Lieferanten und Produktionsprozesse von Waren und Dienstleistungen umfassen, die häufig in verschiedenen Ländern angesiedelt sind. Daher ist es wichtig, statistische Einheiten in globalen Wertschöpfungsketten so weit wie möglich richtig zu identifizieren (vgl. Kapitel 7), um die Kompatibilität mit anderen Datenquellen (wie z. B. Erhebungen über Direktinvestitionen und den Außenhandel) zu verbessern.

Wirtschaftliche Haupttätigkeit

9.25. Unternehmen sollten gemäß der letzten Ausgabe der Internationalen Systematik der Wirtschaftszweige (ISIC Rev. 4) der Vereinten Nationen (VN, 2008) oder einer äquivalenten regionalen/nationalen Klassifikation nach ihrer wirtschaftlichen Haupttätigkeit klassifiziert werden. Die ISIC-Klassifikation erleichtert die internationale Vergleichbarkeit, indem sie die wirtschaftlichen Tätigkeiten der einzelnen Wirtschaftszweige in Abschnitte, Abteilungen, Gruppen

und Klassen gliedert. In den meisten Fällen lässt sich die Grundgesamtheit auf der Ebene des Abschnitts oder der Abteilung bestimmen. Die nachstehenden Empfehlungen beziehen sich auf die Abschnitte und Abteilungen, wie sie in ISIC Rev. 4 definiert sind. Sie sollten mit jeder künftigen Revision der ISIC-Klassifikation aktualisiert werden.

9.26. Wenn sich die tatsächliche wirtschaftliche Tätigkeit von Unternehmen nicht genau bestimmen lässt (z. B. wenn diese Information im Unternehmensregister nicht verfügbar ist, sich auf nicht amtliche Klassifikationen bezieht oder wahrscheinlich veraltet ist), können Innovationserhebungen eine Frage zu den wichtigsten Produktlinien enthalten, die von jedem Unternehmen hergestellt werden, und, wenn möglich, Fragen zur relativen Bedeutung der verschiedenen Arten von Produktlinien (z. B. zum Umsatzanteil verschiedener Produktkategorien). Diese Information ist notwendig, um das Unternehmen für die Zwecke der Schichtenbildung, Stichprobenauswahl und Analyse einer wirtschaftlichen Tätigkeit zuordnen zu können.

9.27. Wie in Kapitel 1 und 2 erörtert, empfiehlt dieses Handbuch – von einigen Ausnahmen abgesehen, die nachstehend erläutert werden – Innovationsdaten für Unternehmen aus den meisten ISIC-Wirtschaftszweigen zu erheben. Wichtige Aspekte, die bei der Festlegung des empfohlenen Erfassungsbereichs von Innovationserhebungen nach wirtschaftlicher Tätigkeit, insbesondere für internationale Vergleichszwecke, berücksichtigt werden sollten, sind der Anteil der nicht unternehmerischen Akteure in einem Wirtschaftszweig, spezifische Messprobleme wie sich ändernde Unternehmensregister und die bisherige internationale Erfahrung mit der Messung von Innovationen in einem Wirtschaftszweig.

9.28. Tabelle 9.1 stellt die allgemeine Wirtschaftszweigstruktur gemäß ISIC Rev. 4 auf Ebene der Abschnitte und Abteilungen dar und erfasst die wirtschaftlichen Tätigkeiten, die sich für internationale Vergleiche empfehlen, die wirtschaftlichen Tätigkeiten, die für nationale Zwecke ergänzend berücksichtigt werden können, und die wirtschaftlichen Tätigkeiten, die sich derzeit nicht für Innovationserhebungen im Unternehmenssektor eignen.

9.29. Die wirtschaftlichen Tätigkeiten, deren Berücksichtigung in nationalen Datenerhebungen und für internationale Vergleiche **empfohlen** wird, umfassen die ISIC Rev. 4 Abschnitte B bis einschließlich M, mit Ausnahme von Abschnitt I (Gastgewerbe). In diesen Bereichen ist eine umfassende nationale und internationale Erfahrung mit der Datenerhebung für Vergleichszwecke vorhanden.

9.30. Die wirtschaftlichen Tätigkeiten, die **ergänzend** erfasst werden sollen, für internationale Vergleichszwecke aber noch weitgehend unerprobt sind, beziehen sich auf die ISIC Rev. 4 Abschnitte A (Land- und Forstwirtschaft, Fischerei), I (Gastgewerbe), N (Erbringung von sonstigen wirtschaftlichen Dienstleistungen) und die Abteilungen 95–96 von Abschnitt S (Reparatur von Datenverarbeitungsgeräten und Gebrauchsgütern, Erbringung von sonstigen überwiegend persönlichen Dienstleistungen). Für diese Wirtschaftszweige ist die internationale Standardisierung der Unternehmensregister noch unvollständig (vor allem für die Landwirtschaft) und die derzeitige Erfahrung beschränkt sich auf Erhebungen in nur wenigen Ländern. Von daher können alle Innovationserhebungen, die diese Wirtschaftszweige adressieren, dazu beitragen, die Empfehlungen für die Messung von Innovationen in diesen Bereichen zu verbessern.

9.31. Für bestimmte wirtschaftliche Tätigkeiten wird *empfohlen, diese generell nicht* in die Datenerfassung im Rahmen von Innovationserhebungen einzubeziehen und bei internationalen Vergleichen zu Innovationen im Unternehmenssektor auszuschließen. Für internationale Vergleichszwecke empfiehlt es sich nicht, Abschnitte O (Öffentliche Verwaltung, Verteidigung; Sozialversicherung), P (Erziehung und Unterricht), Q (Gesundheits- und Sozialwesen), R (Kunst, Unterhaltung und Erholung) und Abteilung 94 von Abschnitt S (Interessenvertretungen sowie kirchliche und sonstige religiöse Vereinigungen [ohne Sozialwesen und Sport]) einzubeziehen,

weil in vielen Ländern staatliche Stellen oder private Organisationen ohne Erwerbszweck bei der Erbringung dieser Dienstleistungen eine dominierende oder wichtige Rolle spielen. Aufgrund nationaler Politikerfordernisse kann es jedoch notwendig sein, den Erfassungsbereich nationaler Erhebungen auf Unternehmen auszudehnen, die in diesen Bereichen tätig sind, z. B. wenn ein erheblicher Anteil der in einem dieser Bereiche tätigen Einheiten im betreffenden Land Unternehmen sind oder wenn diese Unternehmen Anspruch auf öffentliche Förderung für ihre Innovationsaktivitäten haben.

Tabelle 9.1. Gliederung der wirtschaftlichen Tätigkeiten zum Zweck internationaler Vergleiche von Innovationen im Unternehmenssektor

Basierend auf den Abschnitten und Abteilungen der ISIC Rev. 4 der Vereinten Nationen

Abschnitt	Abteilung	Beschreibung
Wirtschaftliche Tätigkeiten, deren Berücksichtigung sich für internationale Vergleiche empfiehlt		
B	05–09	Bergbau und Gewinnung von Steinen und Erden
C	10–33	Verarbeitendes Gewerbe
D	35	Energieversorgung
E	36–39	Wasserversorgung; Abwasser- und Abfallentsorgung und Beseitigung von Umweltverschmutzungen
F	41–43	Baugewerbe
G	45–47	Handel; Instandhaltung und Reparatur von Kraftfahrzeugen
H	49–53	Verkehr und Lagerei
J	58–63	Information und Kommunikation
K	64–66	Erbringung von Finanz- und Versicherungsdienstleistungen
L	68	Grundstücks- und Wohnungswesen
M	69–75	Erbringung von freiberuflichen, wissenschaftlichen und technischen Dienstleistungen
Wirtschaftliche Tätigkeiten, die für die nationale Datenerhebung ergänzend erfasst werden sollen		
A	01–03	Land- und Forstwirtschaft, Fischerei
I	55–56	Gastgewerbe
N	77–82	Erbringung von sonstigen wirtschaftlichen Dienstleistungen
S	95–96	Reparatur von Datenverarbeitungsgeräten und Gebrauchsgütern, Erbringung von sonstigen überwiegend persönlichen Dienstleistungen
Wirtschaftliche Tätigkeiten, die sich für die Datenerhebung nicht empfehlen		
O	84	Öffentliche Verwaltung, Verteidigung; Sozialversicherung
P	85	Erziehung und Unterricht
Q	86–88	Gesundheits- und Sozialwesen
R	90–93	Kunst, Unterhaltung und Erholung
S	94	Interessenvertretungen sowie kirchliche und sonstige religiöse Vereinigungen (ohne Sozialwesen und Sport)
Wirtschaftliche Tätigkeiten, die außerhalb des Erfassungsbereichs dieses Handbuchs liegen		
T	97–98	Private Haushalte mit Hauspersonal; Herstellung von Waren und Erbringung von Dienstleistungen durch Private Haushalte für den Eigenbedarf ohne ausgeprägten Schwerpunkt
U	99	Exterritoriale Organisationen und Körperschaften

9.32. Die anderen Abschnitte, die ausgeklammert werden sollten, werden von Akteuren mit nicht marktbestimmten Tätigkeiten dominiert und liegen daher *außerhalb des Erfassungsbereichs* dieses Handbuchs. Dies betrifft Abschnitt T (Private Haushalte mit Hauspersonal; Herstellung von Waren und Erbringung von Dienstleistungen durch Private Haushalte für den Eigenbedarf ohne ausgeprägten Schwerpunkt) und Abschnitt U (Exterritoriale Organisationen und Körperschaften).

Größe der Einheit

9.33. Obwohl größere Unternehmen in der Regel häufiger eine Innovationstätigkeit aufweisen und ihre Innovationsaktivitäten umfassender sind, können Einheiten jeder Größe innovationsaktiv sein und sollten in die Innovationserhebungen einbezogen werden. Bei kleineren Unternehmenseinheiten, insbesondere jenen mit einem höheren Grad an Informalität (die z. B. nicht als Unternehmen eingetragen sind, von bestimmten Steuern befreit sind oder bestimmte Steuern nicht erklären usw.), ist die Wahrscheinlichkeit jedoch größer, dass sie in statistischen Unternehmensregistern nicht aufgeführt sind. Ihre relative Bedeutung ist u. U. in Ländern höher, die sich in einem früheren Entwicklungsstadium befinden. Vergleiche mit Daten aus Ländern, die Kleinunternehmen getrennt erfassen und in denen der Anteil der Produktion in der informellen Wirtschaft unterschiedlich groß ist, gestalten sich daher als schwierig. Die korrekte Interpretation von Innovationsdaten neu gegründeter Unternehmen, von deren Aktivitäten sehr viele als neu betrachtet werden können, ist eine weitere Herausforderung, wie bereits in Kapitel 3 beschrieben.

9.34. Für internationale Vergleiche wird deshalb empfohlen, den Umfang der Grundgesamtheit auf statistische Unternehmenseinheiten mit *mindestens zehn Beschäftigten* zu beschränken und für die Größenkategorien die *durchschnittliche Kopfzahl* zu verwenden. Je nach Interesse und Ressourcen der Nutzer können in Erhebungen auch Einheiten mit weniger als zehn Beschäftigten berücksichtigt werden, insbesondere in Hochtechnologie- und wissensintensiven Dienstleistungsbranchen. In diese Gruppe dürften Start-ups und Spin-offs fallen, die von erheblichem politischem Interesse sind (vgl. Kapitel 3).

9.2.3. Datenverknüpfung

9.35. Die nationalen Statistikämter nutzen oft das amtliche Unternehmensregister, um die Stichprobe für die Innovationserhebung und die Erhebungen zu FuE, IKT und allgemeinen Unternehmensstatistiken zu bestimmen. Dadurch wird es möglich, die Innovationserhebung mit anderen Erhebungen zu verknüpfen, um metrische Daten zu mehreren interessierenden Variablen wie z. B. zu FuE, IKT, Beschäftigung, Umsatz, Exporten oder Investitionen zu erhalten. Im Lauf der Jahre haben immer mehr nationale Statistikämter Datenverknüpfungen durchgeführt, um teilweise bestimmte Arten von Daten in der Innovationserhebung nicht erfassen zu müssen. Allerdings ist die Datenverknüpfung nur möglich, wenn in den zu verknüpfenden Erhebungen dieselbe statistische Einheit verwendet wird, die für die nationalen Statistikämter in der Regel das Unternehmen ist.

9.36. Die Verknüpfung mit Daten aus verpflichtenden FuE- und anderen Unternehmenserhebungen kann den Aufwand der Antwortpersonen reduzieren, was zu höheren Rücklaufquoten und einer besseren Qualität von metrischen Daten führen kann. Gleichwohl müssen in Innovationserhebungen dann bestimmte Fragen zu allgemeinen Kennzahlen aufgenommen werden, wenn sie als Bezugspunkt für spezielle Fragen benötigt werden und den Antwortpersonen helfen, diese speziellen Fragen in den richtigen Kontext zu stellen, oder wenn sie einen Referenzwert für die Berechnung von Unterkategorien oder Anteilen darstellen. Fragen zu Innovationsaufwendungen sollten beispielsweise als Bezugspunkt eine Frage zu FuE-Aufwendungen umfassen, und Fragen zur Anzahl (oder zum Anteil) der Beschäftigten nach Bildungsniveau sollte eine Frage zur Gesamtzahl der Beschäftigten vorausgehen. Für bestimmte Analysen können die in der Innovationserhebung für FuE, Beschäftigung oder andere Variablen gewonnenen Werte nach Abschluss der Erhebung durch Werte aus den FuE- und Unternehmenserhebungen ersetzt werden, falls die Analyse ergibt, dass dies die Genauigkeit erhöht.

9.37. Eine weitere Möglichkeit, die sich aus der Kombination von Verwaltungs- und Erhebungsdaten ergibt, ist das Vorausfüllen von Online-Innovationsfragebögen mit Angaben aus

anderen Quellen zu Umsatz, Beschäftigung, FuE-Aufwendungen, Patentanmeldungen usw. Dies kann den Antwortpersonen direkte Bezugspunkte liefern und den Beantwortungsaufwand verringern. Ein Nachteil ist, dass vorausgefüllte Daten veraltet sein könnten, wenngleich ältere Daten für das erste Jahr des Beobachtungszeitraums trotzdem nützlich sein können. Die Antwortpersonen sollten außerdem die Möglichkeit haben, Fehler in den vorausgefüllten Feldern zu berichtigen.

9.38. Nach einer ausreichenden Zeitspanne (mindestens ein Jahr nach einer Innovationserhebung) sind Verknüpfungen zwischen strukturellen statistischen Daten von Unternehmen und wirtschaftlichen Variablen für die Forschung nützlich, um Kausalzusammenhänge zwischen Innovationsaktivitäten und -ergebnissen herzustellen. Relevante Ergebnisse sind Veränderungen der Produktivität, der Beschäftigung, der Exporte und der Umsatzzahlen.

9.39. Ausgewählte Innovationsfragen können gegebenenfalls in andere Erhebungen aufgenommen werden, um die Auswahlgesamtheit für Innovationserhebungen zu verbessern, zu aktualisieren oder fortzuschreiben.

9.2.4. Häufigkeit der Datenerhebung

9.40. Wie häufig Innovationserhebungen durchgeführt werden sollen, hängt von praktischen Erwägungen und den Bedürfnissen der Nutzer auf internationaler, nationaler und regionaler Ebene ab. Aspekte wie Kosten, die zeitliche Stabilität von Variablen, die Erhöhung des Beantwortungsaufwands durch häufige Erhebungen und Probleme aufgrund sich überschneidender Beobachtungszeiträume zwischen aufeinanderfolgenden Erhebungen beeinflussen die Häufigkeit, die für die Durchführung von Innovationserhebungen empfohlen wird. Angesichts der Bedeutung von Innovationen für Wirtschaftswachstum und Wohlstand bedarf es für einschlägige Politikmaßnahmen häufig aktualisierter, auf Jahresbasis erhobener Daten, insbesondere für Innovationsaktivitäten, die sich rasch verändern können. Jährliche Panelerhebungen können zudem die Analyse der Lag-Struktur zwischen Innovationsinputs und -outputs oder der Beiträge von Innovationen zur Wirtschaftsleistung erleichtern (vgl. Kapitel 11).

9.41. Es wird empfohlen, alle **ein bis drei Jahre** Innovationserhebungen durchzuführen. Bei einer Häufigkeit von zwei oder drei Jahren kann in den dazwischenliegenden Jahren eine Kurzerhebung vorgenommen werden, die nur die wichtigsten Innovationsvariablen erfasst, sofern die Ressourcen es zulassen. Beim Vergleich der Ergebnisse von Kurz- und Langerhebungen ist jedoch Vorsicht geboten, da die Antworten u. U. durch die Länge der Erhebung beeinflusst werden (vgl. Abschnitt 9.3). In den Jahren, in denen keine Innovationserhebung durchgeführt wird, können Informationen über Innovationen auch aus dem Internet oder anderen Quellen gewonnen werden. Die verschiedenen Möglichkeiten zur Nutzung alternativer Quellen für Innovationsdaten in den Jahren ohne Erhebung müssen noch eingehend untersucht werden.

9.2.5. Beobachtungs- und Referenzzeitraum

9.42. Damit die Antworten der Befragten vergleichbar sind, muss in Erhebungen ein Beobachtungszeitraum für die Fragen zur Innovationstätigkeit angegeben werden. Die meisten, wenngleich nicht notwendigerweise alle Fragen zu Innovationsmerkmalen, sollen sich auf diesen Beobachtungszeitraum beziehen. Um eine Erinnerungsverzerrung bzw. einen Recall-Bias zu verhindern, wird empfohlen, dass der Beobachtungszeitraum **drei Jahre nicht überschreiten sollte**. Der Referenzzeitraum ist das letzte Jahr des gesamten Beobachtungszeitraums einer Erhebung und dient als Beobachtungszeitraum für die Erfassung von metrischen Daten, wie z. B. Aufwendungen oder Anzahl der Beschäftigten. Referenz- und Beobachtungszeitraum sind in Erhebungen identisch, in denen der Beobachtungszeitraum ein Jahr beträgt.

9.43. Die Länge des Beobachtungszeitraums bedingt die Definition von Innovation und damit den Anteil der Einheiten, die als innovativ angegeben werden (vgl. Kapitel 3). So kann beispielsweise die Wahl des Beobachtungszeitraums die Vergleiche zwischen Gruppen von Einheiten (z. B. Wirtschaftszweigen) beeinflussen, die Waren oder Dienstleistungen mit unterschiedlich langen Lebenszyklen produzieren (Wirtschaftszweige mit kurzen Produktlebenszyklen führen mit größerer Wahrscheinlichkeit häufiger Produktinnovationen ein). Dies hat Auswirkungen auf die Interpretierbarkeit und macht eine angemessene Standardisierung von nationalen Erhebungen notwendig (vgl. Kapitel 11).

9.44. In manchen Fällen sprechen Schwierigkeiten bei der Auswertung für einen längeren Beobachtungszeitraum. Wenn ein Innovationsprojekt z. B. über mehrere Jahre läuft, könnte ein kurzer Beobachtungszeitraum dazu führen, dass einzelne Innovationsaktivitäten und -outputs verschiedenen Jahren zugeordnet werden, wie der Abschluss von Kooperationen, der Bezug öffentlicher Fördermittel und der Umsatz mit neuen Produkten. Dies könnte einige relevante Analysen der Innovationsmuster und -auswirkungen beeinträchtigen.

9.45. Bedenken hinsichtlich der Datenqualität sprechen hingegen für einen kürzeren Beobachtungszeitraum, um Erinnerungsfehler zu reduzieren. Diese entstehen z. B., wenn die Antwortpersonen vergessen, ein Ereignis zu melden, oder wenn sie irrtümlich auch Ereignisse aus der Zeit vor dem Beobachtungszeitraum angeben, als wären sie im Beobachtungszeitraum eingetreten (d. h., sie werden in den Beobachtungszeitraum hinein „teleskopiert“).

9.46. Die Qualitätsvorteile bei Daten aus kurzen Beobachtungszeiträumen und die potenziellen Interpretationsvorteile bei Daten aus längeren Beobachtungszeiträumen können in einem Längsschnittpanel kombiniert werden, in dem die Angaben von Unternehmen aus aufeinanderfolgenden Querschnitterhebungen zu Innovationen miteinander verknüpft werden (vgl. Unterabschnitt 9.4.3). Wenn sich die entsprechenden Daten beispielsweise auf einen Beobachtungszeitraum von einem Jahr beziehen, kann der Innovationsstatus von Unternehmen über einen Zeitraum von zwei (drei) Jahren effektiv anhand der Daten für Unternehmen berechnet werden, für die Beobachtungen über zwei (oder drei) aufeinanderfolgende jährliche Beobachtungszeiträume vorliegen. Zusätzliche Annahmen und Anstrengungen sind in Fällen erforderlich, in denen wiederholte Beobachtungen nicht für alle Unternehmen in der Stichprobe verfügbar sind, z. B. aufgrund eines Rückgangs der Unternehmenszahl oder der Verwendung von Stichprobenverfahren, die den Aufwand für bestimmte Gruppen von Befragten reduzieren sollen (z. B. KMU). Für ein Paneldesign spricht das stichhaltige Argument, dass es das Spektrum möglicher Analysen der Kausalzusammenhänge zwischen Innovationsaktivitäten und -ergebnissen erweitert (vgl. Unterabschnitt 9.4.3).

9.47. Beobachtungszeiträume, die länger sind als die Häufigkeit der Datenerhebung, können Vergleiche der Ergebnisse aus aufeinanderfolgenden Erhebungen beeinträchtigen. In solchen Fällen kann es schwierig sein zu bestimmen, ob Veränderungen der Ergebnisse im Zeitverlauf hauptsächlich auf Innovationsaktivitäten im nicht überlappenden Zeitraum zurückzuführen sind oder ob sie durch Aktivitäten im Überschneidungszeitraum mit der vorangegangenen Erhebung beeinflusst sind. Dadurch könnte eine serielle Scheinkorrelation entstehen.

9.48. Zum Zeitpunkt der Veröffentlichung dieses Handbuchs liegt der von den Ländern verwendete Beobachtungszeitraum zwischen einem und drei Jahren. Dies beeinträchtigt die internationale Vergleichbarkeit maßgeblicher Indikatoren wie Verbreitung von Innovationen und Grad der Kollaboration mit anderen Akteuren. Auch wenn es derzeit keinen Konsens über die optimale Länge eines allgemeingültigen Beobachtungszeitraums gibt (abgesehen von einer maximalen Dauer von drei Jahren), würde eine Konvergenz zu einem einheitlichen Beobachtungszeitraum die internationale Vergleichbarkeit erheblich verbessern. Daher wird empfohlen, in gemeinsamen Anstrengungen weitere Ansätze zu den Auswirkungen unterschiedlich langer

Beobachtungszeiträume und zur Verwendung von Paneldaten zu erproben, um Interpretationsschwierigkeiten zu überwinden. Die Ergebnisse dieser Erprobungen könnten dazu beitragen, eine internationale Einigung über die optimale Länge des Beobachtungszeitraums zu erzielen.

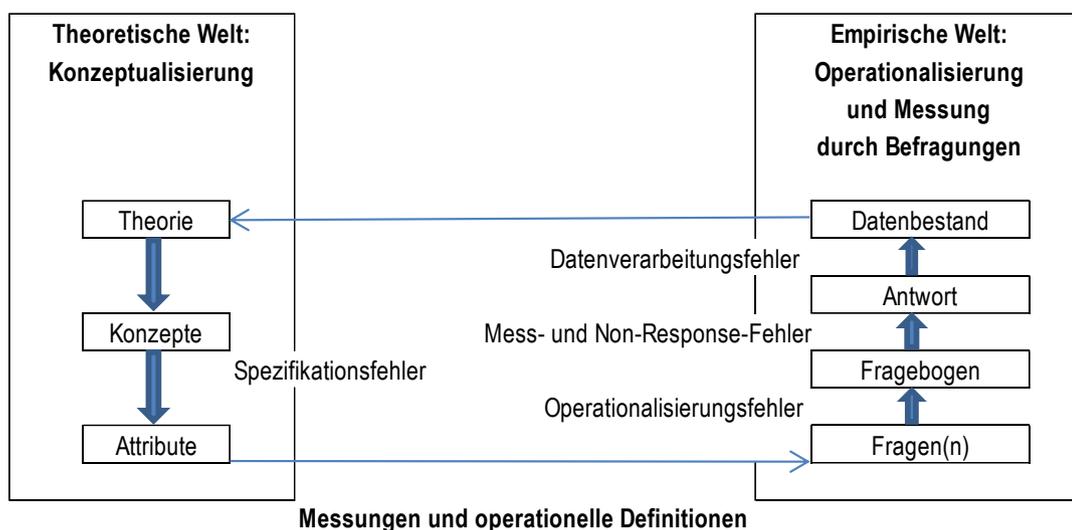
9.3. Fragenformulierung und Fragebogendesign

9.49. In den Kapiteln 3 bis 8 dieses Handbuchs werden verschiedene Konzepte und Merkmale für die Messung von Innovationen im Unternehmenssektor aufgeführt. Diese Konzepte und Merkmale müssen in Form von Fragen operationalisiert werden, die nützliche Daten für die empirische Analyse generieren, wie in Abbildung 9.1 dargelegt.

9.50. Die Operationalisierung von theoretischen Konzepten kann zu einer Reihe von möglichen Fehlern führen, die sich durch ein sorgfältiges Fragen- und Fragebogendesign auf ein Mindestmaß reduzieren lassen. Dieses Handbuch liefert keine vollständigen Beispiele für Erhebungsfragen, da ihre endgültige Formulierung wahrscheinlich vom jeweiligen Kontext abhängt, der zwischen und in den einzelnen Ländern variiert. Stattdessen enthalten die folgenden Abschnitte Leitlinien für empfehlenswerte Praktiken im Hinblick auf das Fragendesign und das Fragebogenlayout. Eine gute Fragenformulierung und ein ansprechendes Fragebogenlayout können die Datenqualität erhöhen, die Rücklaufquoten steigern und die Vergleichbarkeit zwischen verschiedenen Erhebungsmethoden verbessern (vgl. Unterabschnitt 9.5.4).

9.51. Die Formulierung einzelner Fragen und das Layout des Fragebogens hängen davon ab, ob Daten aus anderen Quellen genutzt werden können (was manche Fragen überflüssig macht), wie auch von der Wahl der Erhebungsmethode (vgl. Abschnitt 9.5). Werden beispielsweise Matrix-Fragen verwendet und mündlich gestellt, wie in telefonischen Befragungen, ist ihre Erläuterung schwierig und zeitaufwendig, wohingegen sie einfach zu verstehen sind, wenn sie schriftlich vorgelegt werden, wie beispielsweise in Online- oder postalisch versandten Befragungen. Diese Unterschiede zwischen Erhebungsmethoden müssen berücksichtigt werden, wenn mehrere Erhebungsmethoden zum Einsatz kommen.

Abbildung 9.1. Von der Innovationstheorie zu Innovationsdaten



Quelle: Basierend auf Snijkers und Willimack (2011), „The missing link: From concepts to questions in economic surveys“.

9.3.1. Fragenformulierung

Anpassung und Übersetzung der Fragen

9.52. Alle Fragen müssen sorgfältig angepasst und übersetzt werden, um sicherzustellen, dass die Antwortpersonen die Fragen gemäß den in diesem Handbuch enthaltenen Konzepten und Definitionen interpretieren. Erstens können viele Konzepte und Definitionen nicht unmittelbar in Fragen übertragen werden. Es kann zum Beispiel notwendig sein, mehrere Fragen zu stellen, um Daten zu einem einzigen Konzept zu erheben (vgl. Kapitel 3). Zweitens müssen die Schlüsselbegriffe an die Sprache angepasst werden, die von den Antwortpersonen im jeweiligen kulturellen, regionalen und nationalen Kontext benutzt wird (Harkness et al. [Hrsg.], 2010). In Australien und Kanada wird beispielsweise für ein Unternehmen der Begriff „business“ und nicht der Begriff „enterprise“ verwendet, weil letzterer in beiden Ländern nicht zum allgemeinen englischen Sprachgebrauch gehört und daher missverständlich sein könnte. Die Begriffe „enterprise“ bzw. „business“ könnten auch für Antwortpersonen aus Betrieben oder fachlichen Einheiten verwirrend sein.

9.53. Besonders wichtig sind Übersetzungsfragen bei Innovationserhebungen, die in mehreren Ländern oder Ländern mit mehr als einer Sprache durchgeführt werden, da selbst geringfügige Differenzen zwischen nationalen Fragebögen die Vergleichbarkeit der Ergebnisse verringern können. Diese Differenzen können sich aus der Übersetzung, Veränderungen bei der Anordnung der Fragen oder dem Hinzufügen oder Löschen von Kategorien oder Fragen ergeben. Die Übersetzung muss den landesspezifischen Gegebenheiten Rechnung tragen (wie dem Rechtssystem und den rechtlichen Regelungen eines Landes), um Missverständnisse bei Konzepten und Definitionen zu vermeiden.

Verständlichkeit und Qualität der Fragen

9.54. Die Fragen müssen kurz, in einfacher Sprache und eindeutig formuliert sein. Es ist wichtig, Wiederholungen zu vermeiden, wie zwei Fragen, die auf ähnliche Informationen abzielen, ebenso wie Fragen, die auf mehrere Informationselemente ausgerichtet sind (sie sind oft daran zu erkennen, dass zwei Sätze mit „und“ verbunden sind). Wo immer möglich, sollten die Konzepte und Definitionen in die Fragen aufgenommen werden, da die Antwortpersonen Zusatzinformationen selten lesen. Erläuterungen in Form von Fußnoten oder Hyperlinks bei Online-Fragebögen sollten ebenfalls vermieden werden.

9.55. Die Datenqualität kann verbessert werden, indem versucht wird, die Motivation der Befragten hoch zu halten. Dabei spielt die Länge der Fragen eine wichtige Rolle. Die Motivation kann durch Fragen gesteigert werden, die für die Antwortgebenden relevant und von Interesse sind. Letzteres ist für Befragte aus nicht innovativen Einheiten besonders wichtig: Sie müssen den Fragebogen relevant und interessant finden, andernfalls ist die Wahrscheinlichkeit gering, dass sie ihn beantworten. Im Idealfall sollten alle Fragen für alle Einheiten in allen Wirtschaftszweigen relevant sein (Tourangeau, Rips und Rasinski, 2000).

9.56. „Satisficing“ bezeichnet eine Verhaltensweise von Antwortpersonen, die darauf abzielt, den Aufwand für die Beantwortung eines Online- oder Papierfragebogens möglichst gering zu halten. Antwortpersonen können etwa die Befragung abbrechen, bevor sie vollständig beantwortet ist (vorzeitige Beendigung), Fragen überspringen, für alle Unterfragen einer Frage die gleiche Antwortkategorie wählen (Nichtdifferenzierung) – d. h. beispielsweise auf alle Unterfragen einer Matrixfrage „weniger wichtig“ antworten – oder die Fragen lediglich überfliegen und unpräzise Antworten geben (Barge und Gelbach, 2012; Downes-Le Guin et al., 2012). Die wichtigsten Strategien zur Unterbindung von Satisficing sind, sicherzustellen, dass die Fragen für alle Antwortpersonen von Interesse sind, und den Fragebogen so kurz wie möglich zu halten. Nicht differenzierte Antworten können verringert werden, indem die Zahl der Unterfragen in einer

Matrixfrage auf höchstens sieben begrenzt wird (Couper et al., 2013). Matrixfragen mit mehr als sieben Unterfragen können in mehrere Untergruppen aufgeteilt werden. Eine Matrixfrage mit zehn Unterfragen kann beispielsweise in zwei Themenblöcke mit sechs bzw. vier Unterfragen unterteilt werden.

Nominale und ordinale Antwortkategorien

9.57. Bei qualitativen Fragen können nominale Antwortkategorien („ja“ oder „nein“) oder ordinale Antwortkategorien (Wichtigkeits- oder Häufigkeitsskala) verwendet werden. Nominale Antwortkategorien sind einfach und zuverlässig, liefern aber nur begrenzte Informationen. Ordinale Antwortkategorien können dagegen ein gewisses Maß an Subjektivität mit sich bringen. Beide Arten von Fragen können eine Antwortkategorie „Weiß nicht“ oder „Nicht zutreffend“ erfordern.

9.3.2. Fragebogendesign

9.58. Der Fragebogen sollte so kurz wie möglich und logisch strukturiert sein, und sollte klare Anweisungen enthalten. Bei freiwilligen Erhebungen sinken die Rücklaufquoten (Unit-Response-Quoten, d. h. der Prozentsatz der Stichprobe, der den Fragebogen vollständig beantwortet) mit der Länge des Fragebogens. Die Qualität der Antworten lässt auch nach, je mehr man sich dem Ende des Fragebogens nähert (Galesic und Bosnjak, 2009). Elemente rund um die Durchführung einer Erhebung, die Einfluss auf die Unit-Response-Quoten haben, werden weiter unten in Abschnitt 9.6 erörtert.

9.59. Das Verständnis und die Antwortbereitschaft der Befragten können durch das Fragebogenlayout beeinflusst werden, wobei die empfohlenen Praktiken für Papier- und Online-Fragebögen ähnlich sind. Bei ausgedruckten Fragebögen müssen die Filterregeln bzw. Anleitungen zum Überspringen von Fragen deutlich sichtbar sein. In Bezug auf die Schriftgrößen und den Leerraum auf einer Seite muss das Layout nationalen Präferenzen entsprechen. Wo immer möglich, sollten die Anweisungen wiederholt werden, um die Wahrscheinlichkeit zu erhöhen, dass sie von den Befragten befolgt werden.

Filterfragen

9.60. Filterfragen und Anweisungen zum Überspringen von Fragen leiten die Antwortpersonen je nachdem, wie sie die Filterfragen beantworten, zu verschiedenen Teilen eines Fragebogens weiter. Filterfragen können hilfreich sein, um den Beantwortungsaufwand zu reduzieren, vor allem bei komplexen Fragebögen. Umgekehrt können Filterfragen aber auch einem Satisficing-Verhalten Vorschub leisten, wenn Antwortpersonen eine Filterfrage mit „nein“ beantworten, um sich weitere Fragen zu ersparen.

9.61. Filterfragen und Anweisungen zum Überspringen von Fragen sind weniger notwendig, wenn z. B. Fragen konzipiert werden, die von allen Einheiten unabhängig von ihrem Innovationsstatus beantwortet werden können. So lassen sich zusätzliche Informationen erheben, die für Politikmaßnahmen und Datenanalysen von Nutzen sind. Dennoch sind Filterfragen in bestimmten Situationen notwendig, z. B. wenn eine Reihe von Fragen lediglich für Antwortpersonen relevant ist, die mindestens eine Produktinnovation ausweisen.

9.62. Im Online-Format können Fragen bei Filtersetzung automatisch übersprungen werden. Deshalb wird befürchtet, dass die Beantwortung von Online- und Papierfragebögen zu unterschiedlichen Ergebnissen führt. Im Papierformat können die Antwortpersonen nämlich die übersprungenen Fragen sehen und rückwirkend ihre Antwort bei der Filterfrage ändern, sollten sie zu dem Schluss kommen, dass die übersprungenen Fragen relevant sind. Werden sowohl Online- als auch Papierfragebögen verwendet, können übersprungene Fragen im Online-Fragebogen „ausgegraut“ werden, damit sie für die Antwortgebenden lesbar bleiben. Dies könnte die

Vergleichbarkeit mit der Papierversion verbessern. Wenn in einer Online-Erhebung Paradata – d. h. Daten über das Vorgehen beim Ausfüllen von Erhebungsfragebögen – erfasst werden (vgl. Abschnitt 9.5), kann nachvollzogen werden, wie die Teilnehmenden bei der Beantwortung des Fragebogens vorgehen. So kann festgestellt werden, ob das Ausgrauen einen Effekt auf das Verhalten hat, z. B. ob die Antwortpersonen im Nachhinein eine zuvor gegebene Antwort ändern.

Anordnung der Fragen

9.63. Informationen aus zuvor im Fragebogen gestellten Fragen können einen Einfluss darauf haben, wie eine Frage verstanden wird. Das Hinzufügen oder Löschen einer Frage kann sich daher auf die nachfolgenden Antworten auswirken und die Vergleichbarkeit mit früheren Erhebungen oder Erhebungen in anderen Ländern verringern.

9.64. Fragen zu Aktivitäten, die für alle Einheiten unabhängig von ihrem Innovationsstatus relevant sind, sollten vor den Fragen zur Innovation gestellt werden und dürfen nicht auf die Innovation Bezug nehmen. Dies gilt für etwaige Fragen zu den Unternehmenskapazitäten (vgl. Kapitel 5).

9.65. Wo immer möglich sollten Fragen thematisch angeordnet sein, damit Fragen zu ähnlichen Themen gruppiert sind. So sollten beispielsweise Fragen zu Aktivitäten der Wissensbeschaffung und zu Kollaboration für Innovationszwecke im selben Abschnitt des Fragebogens stehen. Fragen zum Beitrag externer Akteure zu einer bestimmten Innovationsart (Produkt- oder Prozessinnovation) sollten in dem Abschnitt stehen, in dem es um diese spezifische Art von Innovation geht.

9.3.3. Kurzerhebungen

9.66. Für viele kleine Einheiten und Einheiten in Sektoren mit geringer Innovationsaktivität kann der Beantwortungsaufwand für einen umfassenden Innovationsfragebogen im Verhältnis zur Innovationsaktivität hoch sein, was niedrigere Rücklaufquoten zur Folge hat. In solchen Fällen könnten kürzere Fragebögen nützlich sein, die sich auf einen Katalog an Kernfragen konzentrieren. Kurzerhebungen können ebenfalls für Einheiten eingesetzt werden, die in früheren Innovationserhebungen keine Innovationsaktivität angegeben haben. Aus empirischen Untersuchungen für Belgien (Hoskens et al., 2016) und verschiedene Entwicklungsländer (Cirera und Muzi, 2016) geht indessen hervor, dass der Anteil innovativer Unternehmen unter den Teilnehmenden an Kurz- und Langerhebungen merklich abweicht: Bei Kurzerhebungen ist die Quote innovativer Unternehmen höher. Diese Ergebnisse legen den Schluss nahe, dass die Unterschiede bei den Innovationsdaten aus Kurz- und Langerhebungen möglicherweise durch Designfaktoren bedingt sind, die es sorgfältig zu berücksichtigen gilt.

9.3.4. Kombination von Innovationserhebungen mit anderen Unternehmenserhebungen

9.67. Mehrere nationale Statistikämter haben ihre Innovationserhebungen mit anderen Erhebungen im Unternehmenssektor kombiniert. Angesichts der konzeptionellen und empirischen Nähe zwischen FuE und Innovation ist dies insbesondere für FuE-Erhebungen der Fall. Grundsätzlich können verschiedene Arten von Unternehmenserhebungen mit Innovationserhebungen verknüpft werden, z. B. indem Fragen zu Innovationen mit Fragen zu Unternehmensmerkmalen, IKT oder Wissensmanagementpraktiken kombiniert werden.

9.68. Aus der Kombination von Erhebungen ergeben sich mehrere Vorteile:

- Eine kombinierte Erhebung kann den Beantwortungsaufwand für die Berichtseinheiten verringern, sofern es gelingt, doppelte Fragen zu vermeiden, und die kombinierte Erhebung dadurch kürzer und weniger schwierig zu beantworten ist als separate Erhebungen.
- Eine kombinierte Erhebung ermöglicht Analysen des Zusammenhangs zwischen Innovationen und anderen Aktivitäten innerhalb der antwortenden Einheit, wie z. B. der IKT-Nutzung. Dies ist von Vorteil, wenn separate Erhebungen nicht miteinander verknüpft werden können oder wenn die Innovationserhebung und die anderen Erhebungen unterschiedliche Stichproben verwenden.
- Mit einer kombinierten Erhebung lassen sich die Druck- und Portokosten für per Post zugestellte Fragebögen sowie die Follow-up-Kosten für alle Arten von Erhebungen reduzieren.

9.69. Kombinierte Erhebungen weisen jedoch auch Nachteile auf:

- Die Unit- und die Item-Response-Quoten können reduziert werden, wenn der kombinierte Fragebogen wesentlich länger ist als die separaten Fragebögen. Dieses Problem tritt insbesondere bei freiwilligen Erhebungen auf.
- Wenn die Themen hinreichend differenziert sind und sich auf verschiedene Funktionsbereiche innerhalb des Unternehmens beziehen, müssen innerhalb einer Organisation – insbesondere einer großen Organisation – u. U. mehrere Personen kontaktiert werden, um Antworten auf alle Fragen zu bekommen.
- Bei der Kombination einer Innovations- und einer FuE-Erhebung kann es zu Fehlern bei der Interpretation von Fragen zu beiden Themenfeldern kommen. Einige Antwortpersonen aus Einheiten, die keine FuE durchführen, könnten fälschlicherweise davon ausgehen, dass eine Innovation FuE voraussetzt oder dass sie lediglich Innovationen melden sollen, die auf FuE basieren. Wie von manchen Ländern berichtet wurde, die probeweise kombinierte FuE- und Innovationserhebungen durchgeführt haben, kann dies die beobachtete Häufigkeit von Innovationen reduzieren (die beobachtete Häufigkeit von FuE bleibt dagegen unverändert) (z. B. Wilhelmsen, 2012). Zudem könnten manche Antwortpersonen Innovationsaufwendungen irrtümlich als FuE-Aufwendungen ausweisen.
- Die Stichprobenrahmen für die Innovationserhebung und andere Erhebungen im Unternehmenssektor können voneinander abweichen. Wenn Innovations- und FuE-Erhebungen miteinander kombiniert werden, kann die Stichprobe für die Innovationserhebung Wirtschaftszweige (und kleine Einheiten) umfassen, die in FuE-Erhebungen in der Regel nicht berücksichtigt werden.

9.70. Basierend auf den obigen Überlegungen sind bei der Kombination einer Innovationserhebung mit einer oder mehreren anderen Unternehmenserhebungen folgende Leitlinien zu beachten:

- In einer kombinierten FuE- und Innovationserhebung sollte das Risiko einer Konzeptverwechslung durch nicht FuE-durchführende Einheiten verringert werden, indem der Fragebogen in zwei Teile gegliedert und der Teil zu Innovationen vorangestellt wird.
- Wenn Fragen zu Innovationen mit anderen Kategorien von Fragen kombiniert werden, wie z. B. Fragen zu IKT oder Unternehmensmerkmalen, sollten diese Teile getrennt werden. Fragen, die für alle Einheiten relevant sind, sollten den Fragen zur Innovation vorangestellt werden.

- In einer kombinierten FuE- und Innovationserhebung können Probleme, die durch die Konzeptualisierung bedingt sind, noch weiter reduziert werden, indem sichergestellt wird, dass sich der Teil zu FuE nur an Einheiten richtet, bei denen davon ausgegangen wird, dass sie FuE durchführen.
- Um einen Rückgang der Rücklaufquoten zu vermeiden, sollte die Länge einer kombinierten Erhebung mit der kumulierten Länge der separaten Erhebungen vergleichbar sein, vor allem bei freiwilligen Erhebungen.
- Bei Vergleichen der Innovationsergebnisse aus kombinierten Erhebungen mit den Ergebnissen aus separaten Innovationserhebungen ist Vorsicht geboten. Die Erhebungsmethode sollte zudem in allen Einzelheiten dargelegt werden und auch die Schritte beschreiben, die die Gefahr einer Konzeptverwechslung verringern.

9.71. Aufgrund der bereits erwähnten Nachteile, wie der Tatsache, dass bei einigen Antwortpersonen der Eindruck erweckt werden könnte, Innovationen setzen unweigerlich FuE voraus, wird in diesem Handbuch daher generell empfohlen, **FuE- und Innovationserhebungen nicht miteinander zu kombinieren**. Auch wenn dies noch nicht getestet wurde, scheint es zum gegenwärtigen Zeitpunkt weniger problematisch zu sein, eine Innovationserhebung mit anderen Arten von Erhebungen zu kombinieren, wie zum Beispiel Erhebungen zu Geschäftsstrategien oder allgemeinen Unternehmensmerkmalen.

9.3.5. Fragebogen-Testphase

9.72. Innovationserhebungen werden regelmäßig auf den neuesten Stand gebracht, um sie an bekannte Herausforderungen anzupassen und so zu gestalten, dass sie neuen Nutzerbedürfnissen gerecht werden. Es wird nachdrücklich empfohlen, alle neuen Fragen und Layoutmerkmale eines Fragebogens vorab kognitiven Tests zu unterziehen, die im Rahmen von persönlichen Interviews mit Teilnehmenden aus der Auswahlgesamtheit durchgeführt werden.

9.73. Kognitives Testen ist eine von Psycholog*innen und Umfrageforscher*innen entwickelte Methode zur Erfassung verbaler Informationen zu Umfrageantworten. Sie wird eingesetzt, um zu evaluieren, inwieweit sich mit einer Frage (oder Fragengruppe) Konstrukte so messen lassen, wie von den Forschenden beabsichtigt, und ob die Antwortpersonen in der Lage sind, hinreichend präzise Antworten zu geben. Die in kognitiven Interviews gesammelten Erkenntnisse dienen dazu, die Fragen zu verbessern, bevor der Fragebogen an die gesamte Stichprobe geschickt wird (Galindo-Rueda und Van Cruysen, 2016). Nicht erforderlich ist kognitives Testen bei Fragen und Layoutmerkmalen, die bereits vorher getestet wurden, es sei denn, sie wurden in einer anderen Sprache oder einem anderen Land getestet. Eine Darstellung des kognitiven Testens findet sich in Willis (2015, 2005).

9.74. Die Befragten müssen für das kognitive Testen nicht nach dem Zufallsprinzip ausgewählt werden, jedoch sollten mindestens zwei Befragte aus jeder möglichen Kombination der folgenden drei Teilgruppen der Grundgesamtheit gezogen werden: nicht innovative und innovative Einheiten, Einheiten aus den Sektoren Warenproduktion und Dienstleistungen und Einheiten aus den beiden Größenklassen: klein/mittel (10–249 Beschäftigte) und groß (über 250 Beschäftigte). Dies ergibt insgesamt 16 zu befragende Einheiten. Es können zwei (oder mehr) Runden des kognitiven Testens erforderlich sein. In der zweiten Runde werden die nach der ersten Runde an den Fragen vorgenommenen Änderungen getestet.

9.75. Zusätzlich zum kognitiven Testen wird empfohlen, eine Piloterhebung über eine Zufallsstichprobe aus der Grundgesamtheit durchzuführen, wenn an einem Fragebogen wesentliche Änderungen vorgenommen wurden, d. h., wenn z. B. zahlreiche neue Fragen hinzugefügt oder in Kombination mit einer grundlegenden Überarbeitung des Fragebogenlayouts neue Fragen aufgenommen wurden. Piloterhebungen können dazu beitragen, die Anordnung der

Fragen im Fragebogen zu optimieren. Zudem liefern sie nützliche Informationen über Ausfallquoten (Item-Non-Response), logische Inkonsistenzen und die Varianz spezifischer Variablen, was hilfreich ist, um den Stichprobenumfang zu bestimmen (vgl. Unterabschnitt 9.4.2).

9.4. Stichprobenauswahl

9.4.1. Auswahlgesamtheit

9.76. Die in einer Stichproben- oder einer Vollerhebung enthaltenen Einheiten sind aus der Auswahlgesamtheit gezogen. Bei der Erstellung einer Erhebung sollte darauf geachtet werden, dass die beabsichtigte Grundgesamtheit (z. B. alle Unternehmen mit mindestens zehn Beschäftigten) und die Auswahlgesamtheit so weit wie möglich übereinstimmen. In der Praxis kann die Auswahlgesamtheit von der Grundgesamtheit abweichen. Die Auswahlgesamtheit (beispielsweise ein Unternehmensregister) kann Einheiten umfassen, die nicht mehr existieren oder nicht mehr zur Grundgesamtheit gehören. Umgekehrt könnten aufgrund von Verzögerungen bei der Aktualisierung des Registers auch Einheiten fehlen, die zur Grundgesamtheit gehören. So kann es passieren, dass Kleinunternehmen mit raschem Beschäftigungswachstum im Register nicht mit ihrer aktuellen Beschäftigtenzahl erfasst sind.

9.77. Die Auswahlgesamtheit sollte aus dem Referenzjahr der Innovationserhebung stammen. Veränderungen bei den Einheiten während des Referenzzeitraums können Auswirkungen auf die Auswahlgesamtheit haben, insbesondere Umstrukturierungen in der Wirtschaftszweigklassifikation (ISIC-Codes), das Entstehen neuer Einheiten sowie Fusionen, Spaltungen oder Schließungen von Einheiten.

9.78. Zur Erstellung des Stichprobenrahmens stützen sich die nationalen Statistikämter in der Regel auf ein aktualisiertes, amtliches Unternehmensregister, das für statistische Zwecke geführt wird. Andere Organisationen mit Interesse an der Durchführung von Innovationserhebungen haben möglicherweise keinen Zugang zu diesem Unternehmensregister. Als Alternative können privat geführte Unternehmensregister verwendet werden. Diese sind aber häufig weniger aktuell als das amtliche Unternehmensregister und können Fehler beim zugewiesenen Wirtschaftszweig und der Beschäftigtenzahl enthalten. Die Repräsentativität privater Register kann auch geringer sein, wenn die Daten aus Unternehmensbefragungen stammen oder wenn das Register nicht alle Wirtschaftszweige abdeckt. Wird ein Stichprobenrahmen ohne Bezugnahme auf ein amtliches Unternehmensregister erstellt, sollten die Erhebungsfragebögen stets spezifische Fragen zur Überprüfung von Größe und Sektor der antwortenden Einheit umfassen. Einheiten, die die Stichprobenkriterien nicht erfüllen, sollten bei der Prüfung der Datenplausibilität ausgeschlossen werden.

9.4.2. Vollerhebung oder Stichprobenerhebung

9.79. Eine Vollerhebung liefert zwar genauere Daten als eine Stichprobenerhebung, doch ist es generell weder möglich noch wünschenswert, die gesamte Grundgesamtheit zu befragen. Zudem ist eine gut konzipierte Stichprobenerhebung für die Datenerfassung oft effizienter als eine Vollerhebung. Bei einer Stichprobenerhebung sollten stets Zufallsstichproben (mit bekannten Wahrscheinlichkeiten) verwendet werden, um die zu befragenden Einheiten auszuwählen.

9.80. Eine Vollerhebung kann aufgrund gesetzlicher Bestimmungen oder bei einer kleinen Auswahlgesamtheit in einer Stichprobenschicht notwendig sein. In kleinen Ländern oder in bestimmten Sektoren kann eine adäquate Stichprobenziehung für manche Schichten einen Stichprobenumfang ergeben, der dem Umfang der Auswahlgesamtheit nahekommt. In diesem Fall liefert eine Vollerhebung für diese Schicht bessere Ergebnisse bei geringen Zusatzkosten. Eine Vollerhebung kann auch in Schichten eingesetzt werden, die für die Politik besonders relevant sind, wie große Einheiten, die für einen Großteil der FuE-Ausgaben eines Landes

verantwortlich sind, oder besonders wichtige Wirtschaftszweige. Häufig ist es so, dass für die Gruppe der KMU Stichprobenerhebungen und für die Gruppe der Großunternehmen Vollerhebungen durchgeführt werden.

Geschichtete Stichprobenziehung

9.81. Eine einfache Zufallsstichprobe (ein einziger Auswahlatz für alle Stichprobeneinheiten einer Grundgesamtheit) ist eine ineffiziente Methode zur Schätzung des Werts einer Variablen bei einem gewünschten Konfidenzniveau in allen Schichten, da eine große Stichprobe erforderlich ist, um eine ausreichende Stichprobenstärke für jene Schichten zu gewährleisten, die nur wenige Einheiten zählen oder in denen die interessierenden Variablen seltener auftreten. Es ist daher effizienter, für Schichten, die durch die Größe der Einheit und die wirtschaftliche Tätigkeit definiert sind, unterschiedliche Auswahlätze zu verwenden.

9.82. Der optimale Stichprobenumfang für geschichtete Stichprobenerhebungen hängt von dem für die Schätzungen gewünschten Genauigkeitsgrad und dem Ausmaß ab, in dem einzelne Variablen in tabellarischen Ergebnissen kombiniert werden sollen. Der Stichprobenumfang sollte außerdem so gewählt werden, dass der erwarteten Antwortausfallquote, der erwarteten Fehlklassifikationsrate von Einheiten und sonstigen Fehlern in der für die Stichprobenziehung verwendeten Auswahlgesamtheit Rechnung getragen wird.

9.83. Der benötigte Stichprobenumfang kann anhand einer Genauigkeitsgrenze oder eines angestrebten Konfidenzniveaus sowie anhand von Daten über die Anzahl der Einheiten, die Größe der Einheiten und die Variabilität der wichtigsten interessierenden Variablen innerhalb jeder Schicht berechnet werden. Die Varianz der Variablen kann auf der Basis früherer Erhebungen oder für neue Variablen anhand der Ergebnisse einer Piloterhebung geschätzt werden. In der Regel sinkt der erforderliche Auswahlatz mit der Anzahl der Einheiten in der Grundgesamtheit und steigt mit der Größe der Einheiten und der Variabilität des Grundgesamtheitswerts sowie der erwarteten Antwortausfallquote.

9.84. Es wird empfohlen, für heterogene Schichten (in denen die Variabilität der interessierenden Variablen hoch ist) und für kleinere Schichten höhere Auswahlätze zu verwenden. Der Auswahlatz sollte in Schichten mit nur wenigen Einheiten 100 % betragen, z. B. wenn es in einem Wirtschaftszweig oder einer Region nur wenige große Einheiten gibt. Der Größe der Einheiten kann auch durch das pps-Auswahlverfahren (*probability proportional to size*) Rechnung getragen werden, bei dem die Auswahlätze in Schichten mit kleineren Einheiten herabgesetzt werden. Alternativ können die Einheiten jeder Schicht nach Größe oder Umsatz sortiert und systematisch in die Stichprobe einbezogen werden. Für unterschiedliche Schichten können unterschiedliche Stichprobenverfahren verwendet werden.

9.85. Die Schichtung der Grundgesamtheit sollte in Bezug auf die Innovationsaktivitäten möglichst homogene Schichten hervorbringen. Da sich die Innovationsaktivitäten der Einheiten je nach Wirtschaftszweig und Größe der Einheit merklich voneinander unterscheiden, wird empfohlen, die Schichtenbildung anhand der wirtschaftlichen Haupttätigkeit und Größe der Einheiten vorzunehmen. Außerdem kann eine Schichtung nach Region erforderlich sein, um Politikanforderungen gerecht zu werden. Auch eine Schichtung nach Alter einer Einheit kann u. U. notwendig sein.

9.86. Für Größenschichten, die anhand der Beschäftigtenzahl gebildet werden, lautet die Empfehlung:

- kleine Einheiten: 10–49 Beschäftigte
- mittlere Einheiten: 50–249 Beschäftigte
- große Einheiten: über 250 Beschäftigte

9.87. Je nach nationalen Gegebenheiten können auch Schichten für Einheiten mit weniger als 10 und über 500 Beschäftigten gebildet werden. Für eine internationale Vergleichbarkeit ist es jedoch angezeigt, dass die drei oben stehenden Größenschichten abgebildet werden können.

9.88. Die Schichtung der Einheiten nach wirtschaftlicher Haupttätigkeit sollte auf der aktuellen ISIC- oder einer äquivalenten nationalen Wirtschaftszweigklassifikation basieren. Die optimale Gliederungsebene (Abschnitt, Abteilung, Gruppe oder Klasse) hängt weitgehend von den nationalen Umständen ab, die Einfluss auf den für die Erfassung erforderlichen Genauigkeitsgrad haben. Für eine auf die Herstellung von Holzwaren spezialisierte Volkswirtschaft wäre es beispielsweise von Vorteil, für die Tätigkeit (Abschnitt C, Abteilung 16, ISIC Rev 4) eine separate Schicht einzurichten, wohingegen ein Land, dessen Wirtschaft auf den Fremdenverkehr ausgerichtet ist, eher separate Schichten für Abteilung 55 (Beherbergung) und Abteilung 56 (Gastronomie) in Abschnitt I sowie für Abschnitt R (Kunst, Unterhaltung und Erholung) heranziehen würde. Die Stichprobenschichten sollten nicht übermäßig stark aggregiert werden, da dies die Homogenität innerhalb jeder Schicht verringert.

Domänen (Teilgesamtheiten, die von großem Interesse sind)

9.89. Eine Differenzierung der Grundgesamtheit nach Teilgruppen kann für die Datennutzer von besonderem Interesse sein. Es ist zudem möglich, dass Nutzer detaillierte Informationen auf Ebene von Wirtschaftszweigen oder Regionen benötigen. Diese Teilgruppen werden als Domänen (oder Teilgesamtheiten) bezeichnet. Um repräsentative Ergebnisse zu erhalten, muss jede Domäne über die gewählten Stichprobenschichten identifizierbar sein. Am gängigsten ist es, einen hohen Auswahlatz zu verwenden, um zuverlässige Ergebnisse nach Domänen zu erzielen. Durch die Festlegung von Domänen können darüber hinaus verschiedene Unternehmenserhebungen kombiniert und Vergleiche zwischen Einheiten mit ähnlichen Merkmalen im Zeitverlauf vorgenommen werden. Zu den Teilgruppen, die betrachtet werden können, gehören Wirtschaftszweiggruppierungen, Größenklassen, Regionen, in denen Einheiten ansässig sind (Gliedstaat, Provinz, lokale Verwaltungsebene, Kommune, Metropolregion usw.), FuE durchführende Einheiten und Altersklassen. Eine Schichtung nach Alter kann für Studien über junge innovative Unternehmen nützlich sein.

9.90. Relevante Daten nach Domänen können auch außerhalb von repräsentativen Erhebungen der nationalen Statistikämter gewonnen werden, beispielsweise durch Wissenschaftseinrichtungen, Beratungsunternehmen oder andere Organisationen, die Erhebungen oder andere in der Einleitung beschriebene Methoden verwenden. Von Wissenschaftseinrichtungen durchgeführte Erhebungen über Start-ups oder andere Domänen können nützliche Ergebnisse oder nützliche Erfahrungen im Bereich der Datenerfassung liefern, sofern sie geeignete Forschungsmethoden anwenden.

9.4.3. Längsschnittdaten und Querschnittstudien

9.91. Wie zuvor dargelegt, werden Innovationserhebungen in der Regel als aufeinanderfolgende Querschnittstudien durchgeführt, wobei für jede Innovationserhebung eine neue Zufallsstichprobe aus einer gegebenen Grundgesamtheit gezogen wird. Querschnitterhebungen zu Innovationen können in Form eines Längsschnittpanels konzipiert sein, das es ermöglicht, einer Teilmenge von Einheiten über zwei oder mehrere Erhebungswellen einen Katalog identischer Kernfragen zu stellen. Fragen, die nicht als Kernfragen gelten, können von einer Erhebung zur nächsten variieren.

9.92. Paneldaten ermöglichen es, Veränderungen der Innovationsaktivitäten auf mikroökonomischer Ebene im Zeitverlauf zu untersuchen, und erleichtern Forschungsarbeiten, die darauf abzielen, Kausalzusammenhänge zwischen Innovationsaktivitäten und wirtschaftlichen Ergeb-

nissen herzustellen, wie dem Umsatzanteil von Innovationen (vgl. Kapitel 8), wobei die Zeitspanne zwischen einer Innovation und ihren Ergebnissen berücksichtigt werden kann.

9.93. Folgende Vorgehensweisen sollten bei der Erstellung einer Panelerhebung sorgfältig beachtet werden:

- Die Paneleinheiten sollten in die gesamte Querschnittstudie integriert werden, um den Beantwortungsaufwand zu reduzieren, ein akzeptables Maß an Konsistenz zwischen den Ergebnissen der beiden Erhebungen sicherzustellen und qualitativ hochwertige Querschnittsdaten für die Erstellung von Indikatoren zu sammeln. Ein Panel ersetzt nicht die Notwendigkeit einer Querschnittstudie.
- Die Analyse sollte gewährleisten, dass die Einbeziehung der Ergebnisse aus dem Panel die Ergebnisse der Hauptquerschnittstudie weder verzerrt noch anderweitig verfälscht.
- Panelstichproben müssen regelmäßig aktualisiert werden, um Neuzugänge zu erfassen sowie Panelmortalität (Schließung von Einheiten, Abgänge von Einheiten aus der Grundgesamtheit) und Umfragemüdigkeit der Antwortpersonen zu kompensieren. Die Aktualisierung der Stichprobe sollte nach demselben Schichtungsverfahren erfolgen wie für die ursprüngliche Panelstichprobe.

9.5. Erhebungsmethoden

9.94. Es gibt vier Hauptmethoden für die Durchführung von Erhebungen: Online-, postalische, computergestützte telefonische (CATI) oder computergestützte persönliche (CAPI) bzw. persönliche Befragungen. Bei Online- und postalischen Erhebungen lesen die Befragten den Fragebogen selbst. Das Fragebogenlayout spielt daher eine Rolle. CATI- und persönliche Befragungen werden mündlich durchgeführt. Dabei werden den Befragten die Fragen vorgelesen. In einem persönlichen Interview kann dem*der Befragten bei Bedarf jedoch eine gedruckte Version vorgelegt werden.

9.95. In den letzten zehn Jahren sind viele Länder von postalischen zu Online-Erhebungen übergegangen. Die meisten Länder, die als primäre Erhebungsmethode ein Online-Format verwenden, stellen alternativ auch den Fragebogen als herunterladbare Datei (über einen Link in einer E-Mail oder auf der Erhebungswebsite) zur Verfügung oder senden eine gedruckte Version per Post zu.

9.96. Die Wahl der Erhebungsmethode hängt von den Kosten und den möglichen Unterschieden bei den Rücklaufquoten und der Datenqualität ab. Jüngste experimentelle Untersuchungen haben kaum signifikante Unterschiede zwischen Papier- und Online-Datenerhebungen ergeben, weder bei der Antwortqualität noch bei den Rücklaufquoten (Saunders, 2012). Diese Untersuchungen konzentrierten sich aber hauptsächlich auf private Haushalte und evaluierten nur selten Befragungen von Unternehmen. Die Untersuchungen zu den verschiedenen Erhebungsmethoden, insbesondere für den Vergleich mit Online-Formaten, basieren fast ausschließlich auf Befragungen von Studierenden oder Teilnehmenden an kommerziellen Webpanels. Es wäre daher hilfreich, die Wirkung verschiedener Erhebungsmethoden im Unternehmenssektor weiter zu untersuchen.

9.5.1. Postalische Befragungen

9.97. Bei postalischen Befragungen wird den Befragten ein Papierfragebogen zusammen mit einem adressierten und frankierten Rückumschlag zugeschickt, den sie zur Rücksendung des Fragebogens verwenden können. Folgende Vorgehensweise hat sich bewährt: Der*die

Befragte erhält per Post ein Anschreiben und einen Papierfragebogen, dem gegebenenfalls zwei oder drei auf dem Postweg zugestellte Erinnerungsschreiben an nicht antwortende Einheiten und bei Bedarf auch noch telefonische Erinnerungsanrufe folgen sollten.

9.98. Postalische Befragungen ermöglichen es den Antwortpersonen, den gesamten Fragebogen auf einfache Weise rasch anzusehen, um so die Länge, die behandelten Themen und die Relevanz einzuschätzen. Die Beantwortung eines Fragebogens in Papierform kann bei Bedarf problemlos zwischen mehreren Antwortpersonen aufgeteilt werden, beispielsweise wenn die Fragen im Abschnitt zu Innovationsaufwendungen nur von einer Fachkraft aus der Buchhaltung beantwortet werden können (vgl. Abschnitt 9.6 zu mehreren Befragten). Ein gedruckter Fragebogen mit Filterfragen setzt voraus, dass die Antwortpersonen die Anweisungen zur Reihenfolge der Fragen genau befolgen.

9.5.2. Online-Befragungen

9.99. Die bewährte Vorgehensweise bei Online-Befragungen besteht darin, ein einleitendes Schreiben per Post zu versenden, in dem der Zweck der Erhebung erläutert wird, gefolgt von einer E-Mail, die einen Hyperlink zur Erhebung enthält. Für den Zugang sollten eine sichere Kennung und ein sicheres Passwort erforderlich sein, und es sollten aktuelle IT-Sicherheitsmethoden verwendet werden. Das Follow-up umfasst zwei oder drei per E-Mail oder Post zugestellte Erinnerungsschreiben an nicht antwortende Einheiten und, falls erforderlich, telefonische Erinnerungsanrufe.

9.100. Online-Fragebögen können, falls notwendig, zwischen mehreren Befragten aufgeteilt werden, sofern der*die ursprüngliche Adressat*in den anderen Befragten Benutzernamen und Passwort mitteilt (vgl. Abschnitt 9.6).

9.101. Online-Erhebungen haben gegenüber postalischen Erhebungen mehrere Vorteile in Bezug auf die Datenqualität und die Kosten:

- Softwareprogramme können die Antwortpersonen über ein Pop-up-Fenster benachrichtigen, wenn eine Frage nicht beantwortet wurde oder einen Fehler enthält, z. B. wenn ein Wert den erwarteten Höchstwert übersteigt oder wenn ein Prozentsatz 100 % überschreitet. Im Fall einer postalischen Erhebung müssen die Antwortpersonen dagegen telefonisch kontaktiert werden, falls Fehler zu korrigieren sind. Diese Anrufe erfolgen möglicherweise erst mehrere Wochen nach dem Ausfüllen des Fragebogens. Aufgrund der hohen Follow-up-Kosten passiert es bisweilen, dass fehlende Werte bei einer postalischen Erhebung nachträglich durch Imputation ersetzt werden.
- Pop-up-Textfelder, die neben der jeweiligen Frage platziert sind, können verwendet werden, um zusätzliche Informationen anzubieten, wenngleich die Befragten diese Funktion selten nutzen.
- Anders als beim Papierfragebogen können die Befragten bei einer Online-Befragung nicht unmittelbar alle Fragen sehen. Entsprechend ist die Gefahr geringer, dass sie die Antwortoption „Nein“ wählen, um die Beantwortung von Folgefragen zu vermeiden. Eine Online-Befragung kann somit die Zahl von falschen negativen Antworten verringern.
- Die Erhebungskosten sind im Vergleich zu anderen Erhebungsmethoden aus mehreren Gründen niedriger. So müssen die Antwortgebenden bei manchen Arten von Fehlern nicht mehr zwecks Korrektur kontaktiert werden, die Daten werden automatisch in einer Datei gespeichert, die Kriterien für die Prüfung der Datenplausibilität sind weniger streng als bei anderen Methoden und die Kosten für Versand und Druck sind niedriger.

9.102. Der größte Nachteil einer Online-Erhebung im Vergleich zu anderen Erhebungsmethoden ist, dass manche Antwortpersonen u. U. nicht in der Lage oder nicht bereit sind, einen Online-Fragebogen auszufüllen. In diesem Fall bedarf es einer alternativen Erhebungsmethode (vgl. Unterabschnitt 9.5.4). Das Online-System muss u. U. so konzipiert sein, dass verschiedene Personen innerhalb einer Einheit einzelne Erhebungsabschnitte beantworten können.

Erhebung von Paradata bei Online-Befragungen

9.103. Online-Umfrage-Tools bieten die Möglichkeit, anhand von Tastatureingaben und Mausclicks Paradata zu erheben (z. B. um festzustellen, ob auf Hilfsmenüs zugegriffen wurde) oder die Zeit zu ermitteln, die zur Beantwortung spezifischer Fragen, Abschnitte oder der gesamten Erhebung benötigt wurde (Olson und Parkhurst, 2013). Paradata können genutzt werden, um zu ermitteln, mit welchen Methoden sich ein unerwünschtes Verhalten der Antwortpersonen, wie ein vorzeitiger Abbruch der Befragung oder Satisficing, verhindern lässt, sowie um nachzuvollziehen, welche Fragen für die Antwortpersonen schwer verständlich sind (z. B. wenn die Beantwortungszeit für eine Frage deutlich länger ist als im Durchschnitt für eine Frage ähnlicher Art). Sie können auch dabei helfen zu ermitteln, ob Spätantwortende anders als Frühantwortende den Fragebogen deutlich rascher beantworten, wodurch die Datenqualität sinken kann (Belfo und Sousa, 2011; Fan und Yan, 2010; Revilla und Ochoa, 2015).

9.104. Es wird empfohlen, **bei Online-Befragungen Paradata zu erheben**, um Probleme im Zusammenhang mit der Fragenformulierung und dem Fragebogenlayout zu identifizieren.

9.5.3. Telefonische und persönliche Befragungen

9.105. Telefonische und persönliche Befragungen werden mit computergestützten Datenerfassungssystemen durchgeführt. Bei beiden Methoden werden die Fragen den Antwortpersonen vorgelesen, was im Vergleich zu visuellen Erhebungsmethoden, bei denen die Befragten die Fragen vor sich sehen, Veränderungen der Frageformate erforderlich machen kann. Die Interviewer*innen müssen in Befragungstechniken und der Beantwortung von Fragen der Antwortpersonen geschult werden, damit die Antworten der Antwortpersonen nicht durch Interaktionen mit dem*der Interviewer*in verzerrt werden. Filterfragen können in beiden Formaten automatisch zum Einsatz kommen, und die Antwortpersonen hören übersprungene Fragen nicht. Allerdings kann der*die Interviewer*in zusätzliche Fragen stellen, um sicherzugehen, dass z. B. eine gegebene „Nein“- oder „Ja“-Antwort wirklich zutreffend ist.

9.106. Die CATI-Erhebungsmethode bietet im Vergleich zu anderen Methoden eine Zeitersparnis, da die Ergebnisse innerhalb weniger Wochen vorliegen können. Wie Online-Erhebungen können sowohl CATI- als auch CAPI-Befragungen die Zahl der Fehler und fehlenden Werte reduzieren. Ihr größter Nachteil im Vergleich zu einer Online-Erhebung besteht in den höheren Kosten aufgrund des Bedarfs an geschulten Interviewer*innen. Außerdem sind die CATI- und CAPI-Erhebungsmethoden im Vergleich zu Online- und postalischen Befragungen nicht für die Erhebung quantitativer Daten geeignet, sofern die Antwortpersonen entsprechende Informationen erst aufwendig recherchieren müssen.

9.107. Das CAPI-Format wird hauptsächlich verwendet, um hohe Rücklaufquoten zu erzielen. Dies kann z. B. in Kulturkreisen angezeigt sein, in denen persönliche Interviews notwendig sind, um dem*der Befragten Respekt zu zollen, sowie in Bereichen, in denen Online- oder postalische Befragungen unzuverlässig sind.

9.5.4. Kombinierte Erhebungsmethoden

9.108. Die Verwendung mehrerer Erhebungsmethoden kann die Rücklaufquoten deutlich erhöhen (Millar und Dillman, 2011). Wenn möglich, sollten bei Erhebungen komplementäre Erhebungsmethoden kombiniert werden, d. h. visuelle (Papier- oder Online-Fragebogen) oder

auditive (CATI- oder persönliches Interview) Methoden, da die Antworten der Befragten je nach Erhebungsmethode unterschiedlich ausfallen. Bei sozial erwünschten Fragen können telefonische Befragungen ebenfalls höhere Rücklaufquoten liefern als Online- oder postalische Erhebungen (Zhang et al., 2017). Da Innovationen als sozial erwünscht gelten, könnte dies dazu führen, dass bei CATI-Befragungen höhere Innovationsraten angegeben werden als bei Papier- oder Online-Befragungen. Bei der Erstellung von Indikatoren und dem Vergleich von Ergebnissen zwischen Ländern, die verschiedene Erhebungsmethoden anwenden, sollte möglichen Effekten der Erhebungsmethoden Rechnung getragen werden.

9.6. Erhebungskonzept

9.109. Im Erhebungskonzept sind alle Etappen der Erhebungsdurchführung festgehalten, von der Kontaktaufnahme mit den Antwortpersonen, über die Annahme der ausgefüllten Fragebögen bis zum Nachfassen bei nicht antwortenden Einheiten. Das Protokoll sollte im Voraus festgelegt werden und so konzipiert sein, dass alle Befragten die gleiche Chance haben, den Fragebogen zu beantworten, da das Ziel darin besteht, die Rücklaufquote zu maximieren. Es ist jedoch anzunehmen, dass das optimale Erhebungskonzept von Land zu Land variiert.

9.6.1. Identifizierung der Antwortpersonen

9.110. Bei Innovationserhebungen ist die Wahl der geeigneten Antwortpersonen (oder der geeigneten Abteilung innerhalb eines großen Unternehmens) besonders wichtig, weil spezielle Fragen gestellt werden, die nur von wenigen Personen beantwortet werden können. Selten handelt es sich dabei um die Personen, die andere statistische Fragebögen ausfüllen. In kleinen Einheiten sind Geschäftsführer*innen oft geeignete Antwortpersonen. So weit wie möglich sollten sachkundige Antwortpersonen ausgewählt werden, um das physische oder virtuelle Weiterleiten eines Fragebogens an verschiedene Personen im Unternehmen auf ein Minimum zu reduzieren. Werden Fragebögen weitergeleitet, steigt die Wahrscheinlichkeit, dass sie verloren gehen, verlegt werden oder sich niemand dafür zuständig fühlt. In großen Einheiten, in denen wahrscheinlich keine Person allein alle Fragen beantworten kann, lässt sich ein Weiterleiten meist nicht vermeiden. Allerdings sollte eine einzige Kontaktperson oder eine Abteilung bestimmt werden, die für die Koordinierung der Beantwortung des Fragebogens verantwortlich ist.

9.6.2. Erläuterungen für Antwortpersonen

9.111. Innovationserhebungen enthalten Begriffe und Fragen, die u. U. nicht von allen Personen eindeutig verstanden werden. Die Erhebungsleiter*innen sollten das in der Erhebung eingesetzte Personal schulen, damit es etwaige Rückfragen beantworten kann, und ihm eine Liste mit grundlegenden Definitionen und Erläuterungen zu den Fragen zur Verfügung stellen.

9.6.3. Verpflichtende und freiwillige Erhebungen

9.112. Die Teilnahme an Innovationserhebungen kann entweder freiwillig oder verpflichtend sein. Bei einer Pflichtbeteiligung gibt es unterschiedliche Durchsetzungsgrade. Bei freiwilligen Erhebungen wird mit höheren Antwortausfallquoten gerechnet, die mit der Länge des Fragebogens zunehmen. Um erwarteten Antwortausfallquoten Rechnung zu tragen, können die Auswahlsätze erhöht werden. Mögliche Verzerrungen infolge von Unterschieden bei den Besonderheiten nicht antwortender und antwortender Einheiten, die mit den Erhebungsfragen korreliert sind, werden dadurch aber nicht beseitigt. Um Verzerrungen gering zu halten, sollten die Rücklaufquoten und die Repräsentativität maximiert werden (siehe unten).

9.113. Ob eine Erhebung freiwillig oder verpflichtend ist, kann ebenfalls Einfluss auf die Ergebnisse haben. In einer freiwilligen Erhebung kann beispielsweise der berechnete Anteil

innovativer Unternehmen nach oben verzerrt sein, falls Manager*innen nicht innovativer Unternehmen mit geringerer Wahrscheinlichkeit antworten sollten als Manager*innen innovativer Unternehmen (Wilhelmsen, 2012).

9.6.4. Antwortausfall

9.114. Eine Nicht-Teilnahme (Unit-Non-Response) liegt vor, wenn eine Stichprobeneinheit nicht an einer Erhebung teilnimmt. Dieser Fall kann eintreten, wenn das Erhebungsinstitut die Berichtseinheit nicht erreichen kann oder diese die Teilnahme verweigert. Item-Non-Response bezieht sich auf die Antwortausfallquote bei einer bestimmten Frage und entspricht dem Anteil der fehlenden Antworten an allen antwortenden Einheiten. Die Item-Non-Response-Quoten sind bei quantitativen Fragen oft höher als bei Fragen mit nominalen oder ordinalen Antwortkategorien.

9.115. Unit- und Item-Non-Response sind nur wenig problematisch, wenn die fehlenden Antworten nach dem Zufallsprinzip auf alle Stichprobeneinheiten und alle Fragen verteilt sind. Wenn Unit-Non-Response zufallsbedingt ist, kann die statistische Aussagekraft durch Erhöhung des Auswahlsatzes aufrechterhalten werden. Wenn Item-Non-Response zufallsbedingt ist, können einfache Gewichtungsmethoden verwendet werden, um den Wert einer Variablen für die Grundgesamtheit zu schätzen. Beide Arten von Antwortausfällen können jedoch Verzerrungen unterliegen. Beispielsweise können Manager*innen nicht innovativer Einheiten weniger geneigt sein, zu antworten, weil sie den Fragebogen für wenig relevant halten, was dazu führt, dass der Anteil innovativer Einheiten in der Grundgesamtheit überzeichnet ist. Es kann aber auch sein, dass Manager*innen innovativer Einheiten aus Zeitgründen weniger bereit sind, zu antworten.

Erhöhung der Rücklaufquoten

9.116. Hohe Rücklaufquoten lassen sich insbesondere bei freiwilligen Erhebungen einfacher erzielen, wenn die Fragen gut formuliert sind, das Fragebogendesign ansprechend ist und die Erhebungskonzepte sorgfältig konzipiert sind (vgl. Abschnitt 9.3). Zwei Aspekte des Erhebungskonzepts können einen stark positiven Effekt auf die Rücklaufquoten haben: 1. ein gutes Follow-up mit mehrfachen Erinnerungen an nicht antwortende Einheiten und 2. die Personalisierung aller Kontakte, indem in Erinnerungsmails die Antwortpersonen namentlich angesprochen werden und der Wortlaut der einzelnen Erinnerungsmails geändert wird. Zur Personalisierung gehört auch eine erste schriftliche Kontaktaufnahme per Post, die die Rücklaufquoten im Vergleich zu einer ersten Kontaktaufnahme per E-Mail deutlich erhöhen kann (Dykema et al., 2013). Eine klare Kommunikation des Verwendungszwecks der Erhebungsdaten ist für die Schaffung eines Vertrauensklimas und die Teilnahmebereitschaft von entscheidender Bedeutung. Die Beteiligung kann weiter erhöht werden, wenn sich die Manager*innen von wahrheitsgetreuen und sorgfältig überlegten Antworten direkte Vorteile für ihr Unternehmen versprechen.

Umgang mit niedrigen Unit-Response-Quoten

9.117. Es gibt keine klaren Abgrenzungen für hohe, mittlere und niedrige Unit-Response-Quoten. Als Faustregel gilt, dass hohe Rücklaufquoten 70 % bzw. 80 % übersteigen, mittlere Rücklaufquoten bei 50–70 % bzw. 80 % liegen und niedrige Rücklaufquoten weniger als 50 % betragen.

9.118. Außer im Fall einer sehr hohen Rücklaufquote (mehr als 95 %), sollten die Unterschiede zwischen antwortenden und nicht antwortenden Einheiten anhand von Schichtungsvariablen wie Größe der Einheit oder Wirtschaftszweig verglichen werden. Wenn die Rücklaufquote hoch ist und die Verwendung von Schichtungsvariablen keine signifikanten Unterschiede ergibt, kann die Gewichtung der Grundgesamtheit auf der Grundlage der antwortenden Ein-

heiten berechnet werden. Dieses Verfahren beruht auf der Annahme, dass das Innovationsverhalten von antwortenden und nicht antwortenden Einheiten in Bezug auf diese Merkmale identisch ist. Herausforderungen können sich ergeben, wenn das Verhalten innerhalb einer Schicht sehr heterogen ist (z. B. zwischen großen und sehr großen Unternehmen).

9.119. Bei mittleren oder niedrigen Rücklaufquoten wird empfohlen, eine Nicht-Teilnehmer-Befragung durchzuführen (vgl. Unterabschnitt 9.6.5).

9.120. Wenn die Unit-Response-Quote sehr niedrig ist (weniger als 20 %), reicht eine Nicht-Teilnehmer-Befragung möglicherweise nicht aus, um mögliche Verzerrungen zu korrigieren, es sei denn, sie ist von sehr hoher Qualität und erfasst einen Großteil der nicht antwortenden Einheiten. Die Daten können analysiert werden, um festzustellen, ob die Rücklaufquoten in bestimmten Schichten akzeptabel sind, und um für diese Schichten eine Nicht-Teilnehmer-Befragung durchzuführen. Ansonsten sollten die Ergebnisse nicht zur Schätzung der Merkmale der Grundgesamtheit verwendet werden, da die Wahrscheinlichkeit verzerrter Ergebnisse hoch ist. Die Daten können genutzt werden, um Korrelationsstrukturen zwischen Variablen zu untersuchen, solange die Ergebnisse nicht auf die zugrunde liegende Grundgesamtheit verallgemeinert werden.

9.6.5. Durchführung von Nicht-Teilnehmer-Befragungen

9.121. Viele nationale Statistikämter haben ihre eigenen Vorgaben, wann eine Nicht-Teilnehmer-Befragung notwendig ist. Falls keine solche Vorgabe existiert, wird die Durchführung einer Nicht-Teilnehmer-Befragung empfohlen, sobald die Ausfallquote (Unit-Non-Response) in einer Schicht 30 % übersteigt. Die Nicht-Teilnehmer-Befragung sollte mindestens 10 % der nicht antwortenden Einheiten erfassen (mehr bei kurzen Erhebungen oder Schichten mit geringer Grundgesamtheit).

9.122. Ziel der Nicht-Teilnehmer-Befragung ist es, mögliche signifikante Unterschiede bei den Innovationsaktivitäten antwortender und nicht antwortender Einheiten zu ermitteln. Um künftige Erhebungen zu verbessern, können Informationen darüber gewonnen werden, warum sich die nicht antwortenden Einheiten nicht beteiligt haben. Im Idealfall ist die Unit-Response-Quote in der Nicht-Teilnehmer-Befragung ausreichend hoch und sind die Antworten hinreichend zuverlässig, um die Anpassung der Gewichtungen der Grundgesamtheit zu ermöglichen. Allerdings sollten bei der Anpassung der Gewichtungen auch die Effekte der in der Nicht-Teilnehmer-Befragung verwendeten Erhebungsmethode (andere Erhebungsmethoden oder andere Fragebögen als in der Haupterhebung) berücksichtigt werden.

9.123. Der Fragebogen einer Nicht-Teilnehmer-Befragung muss kurz sein (nicht länger als eine Druckseite), und das Ausfüllen darf nicht mehr als zwei oder drei Minuten in Anspruch nehmen. Die wichtigsten Fragen müssen Wort für Wort die „Ja oder Nein“-Fragen der Haupterhebung zu den Innovationsoutputs (Produktinnovationen und Prozessinnovationen) und bestimmten Innovationsaktivitäten (z. B. FuE, Aktivitäten im Bereich Konstruktion, Design und sonstige kreative Arbeit usw.) beinhalten. Gehen aus anderen Quellen keine Informationen zur wirtschaftlichen Tätigkeit und Größe der Berichtseinheit hervor, muss die Nicht-Teilnehmer-Befragung entsprechende Fragen enthalten.

9.124. Nicht-Teilnehmer-Befragungen werden in der Regel in Form von computergestützten telefonischen Befragungen durchgeführt. CATI-Befragungen bieten den Vorteil einer raschen Umsetzung und ermöglichen zugleich hohe Rücklaufquoten für einen kurzen Fragebogen, sofern alle Unternehmen in der Stichprobe telefonisch kontaktiert werden können. Der Nachteil dieser Methode im Vergleich zu einer postalischen oder Online-Erhebung besteht darin, dass in einigen Ländern kurze telefonische Befragungen mehr positive Antworten auf Fragen zu Innovationsaktivitäten und -outputs auslösen als die ursprüngliche Erhebung. Die diesbezüglichen

Erfahrungen sind gemischt, was in unterschiedlichen Ergebnissen in verschiedenen Ländern zum Ausdruck kommt. Es wird empfohlen, weitere experimentelle Untersuchungen zur Vergleichbarkeit der Methoden für Erhebungen im Unternehmenssektor durchzuführen.

9.7. Datenverarbeitung nach der Erhebung

9.125. Die Datenaufbereitung umfasst Fehlerprüfungen, die Imputation fehlender Werte und die Berechnung von Gewichten (Hochrechnungsfaktoren).

9.7.1. Fehlerprüfungen

9.126. Wie in den Unterabschnitten 9.5.2 und 9.5.3 weiter oben erörtert, können durch den Einsatz von Online-, CATI- und CAPI-Erhebungsmethoden potenzielle Fehler automatisch erkannt und die Antwortpersonen unmittelbar um Korrektur gebeten werden. Bei Papierfragebögen müssen alle nachstehenden Arten von Fehlerprüfungen durchgeführt werden, bei Online-Erhebungen ist u. U. jedoch nur die Prüfung der außerhalb des Erfassungsbereichs liegenden Einheiten erforderlich. Wenn Fehler festgestellt werden, sollte die Antwortperson oder die Berichtseinheit so bald wie möglich kontaktiert und um Korrektur gebeten werden.

Außerhalb des Erfassungsbereichs liegende Einheiten

9.127. Bisweilen gehen auch Antworten von Einheiten ein, die außerhalb des Erfassungsbereichs liegen und nicht der festgelegten Grundgesamtheit angehören. Hierzu zählen z. B. Einheiten, deren Beschäftigtenzahl unter der Mindestbeschäftigtenzahl liegt, Einheiten, die sich nicht im Besitz eines Unternehmens befinden, oder Einheiten in einer nicht berücksichtigten ISIC-Kategorie. Die Antworten dieser Einheiten müssen von weiteren Analysen ausgeschlossen werden.

Prüfungen der Datenvalidierung

9.128. Bei diesen Verfahren wird geprüft, ob die Antworten zulässig sind. Die zulässigen Werte für einen prozentuellen Anteilswert liegen z. B. zwischen 0 und 100.

9.129. Eine zusätzliche Prüfung der Datenqualität sollte bei Verhältniszahlen und metrischen Daten vorgenommen werden, vor allem bei Innovationsaufwendungen. Die bewährte Vorgehensweise besteht derzeit darin, metrische Daten mit Daten aus anderen verfügbaren Quellen zu vergleichen (z. B. zu FuE-Aufwendungen und Investitionsausgaben). Darüber hinaus sollten die Schätzwerte des Umsatzanteils von Innovationen und andere metrische Daten auf Ausreißer oder sonstige unerwartete Werte geprüft werden. Diese Methoden sind besonders wichtig für große Einheiten, auf die ein hoher Anteil der insgesamt ausgewiesenen FuE- und Innovationsaufwendungen entfällt.

Prüfung der Relation

9.130. Bei diesen Kontrollen wird die Relation zwischen zwei Variablen evaluiert. Zudem können Hard und Soft Errors identifiziert werden. Von Hard Errors ist die Rede, wenn eine Relation offensichtlich falsch ist, z. B. wenn die Summe von prozentuellen Anteilswerten nicht 100 % ergibt oder wenn die gemeldete Anzahl der Beschäftigten mit Tertiärabschluss über der gemeldeten Gesamtzahl der Beschäftigten liegt. Durch die Prüfung der Relation lassen sich auch Soft Errors ermitteln, bei denen eine Antwort falsch sein könnte. Weist z. B. eine Einheit mit zehn Beschäftigten Innovationsaufwendungen in Höhe von 10 Mio. EUR aus, ist das möglich, aber unwahrscheinlich.

Prüfung auf Fehler bei der Filterführung

9.131. Hierbei wird geprüft, ob alle Fragen, die beantwortet werden sollten, beantwortet wurden, z. B. ob die Befragten aus innovationsaktiven Einheiten alle Fragen zu Innovationsaufwendungen beantwortet haben. Fehler deuten darauf hin, dass die Antwortperson die Filteranweisungen nicht verstanden oder befolgt hat.

9.7.2. Imputation fehlender Daten

9.132. Eine andere Art von Fehler liegt vor, wenn eine Antwortperson eine Frage nicht beantwortet und beispielsweise die Zellen zu mehreren Unterfragen in einer Matrixfrage entweder absichtlich oder versehentlich leer lässt. Die Befragten können sich auch weigern, eine Frage zu beantworten, wenn sie der Meinung sind, dass keine der Antwortkategorien zutrifft, oder wenn sie die Antwort nicht kennen und die Option „Weiß nicht“ nicht vorgesehen ist. In Online-Umfragen können die Befragten zu einer Antwort gezwungen werden. Dies wird bei freiwilligen Erhebungen jedoch nicht empfohlen, da es die Befragten dazu veranlassen kann, die Teilnahme abzubrechen.

9.133. Um die Kosten und den Beantwortungsaufwand zu reduzieren, können fehlende Werte in manchen Fällen anhand zusätzlicher Informationen ermittelt werden, anstatt die Befragten erneut zu kontaktieren. Die Einbeziehung zusätzlicher Informationen sollte eine genauere Schätzung der fehlenden Werte ermöglichen als die bloße Verwendung des beobachteten Mittelwerts einer Schicht. Imputierte Werte sollten stets gekennzeichnet werden, damit sie in multivariaten Analysen nicht als abhängige Variablen verwendet werden und so zu verzerrten Ergebnissen führen.

9.134. Bei Cold-Deck-Imputationsverfahren werden die fehlenden Werte anhand von Daten aus anderen statistischen Erhebungen (einschließlich früherer Erhebungen) oder aus anderen einschlägigen Quellen geschätzt. Daten zur Anzahl der Beschäftigten mit Tertiärabschluss könnten beispielsweise aus Erhebungen entnommen werden, die bei ähnlichen Arten von Einheiten durchgeführt wurden.

9.135. Bei Hot-Deck-Imputationen werden andere Daten aus der Innovationserhebung verwendet, um fehlende Werte zu berechnen. Welches Hot-Deck-Imputationsverfahren konkret gewählt wird, hängt vom Messniveau der Variablen ab. Metrische Daten können entweder anhand des Mittelwerts der Variablen in der Schicht der antwortenden Einheit oder mittels einer Regression berechnet werden, um den Wert der metrischen Variablen vorherzusagen. Im letzteren Fall müssen die Ergebnisse überprüft werden, um unplausible Schätzwerte, wie z. B. negative Werte, auszuschließen.

9.136. Zur Bestimmung fehlender Ausprägungen bei Nominal- und Ordinalskalen können Nearest-Neighbour-Verfahren eingesetzt werden. Bei dieser Methode werden Daten aus fehlerfreien Datensätzen verwendet (Spenderdatensätze, die allen Kriterien der Fehlerprüfungen gerecht werden), um den fehlenden Wert durch den Wert des Spenderdatensatzes zu ersetzen. Dafür werden Spender ausgewählt, die ähnliche Schichtungsvariablen aufweisen und im Hinblick auf verwandte Variablen maximal vergleichbar sind. Um beispielsweise eine fehlende ordinale Variable in Bezug auf Kollaborationspartner zu ergänzen, sollte ein Spenderdatensatz gewählt werden, bei dem verwandte Variablen in Bezug auf Informationsquellen möglichst ähnlich sind.

9.7.3. Berechnung von Gewichten

9.137. Die Ergebnisse von Stichprobenerhebungen müssen gewichtet werden, um repräsentative Schätzwerte für die Grundgesamtheit zu liefern. Für die Gewichtung der Stichprobenergebnisse gibt es verschiedene Methoden. Die einfachste Methode ist die Gewichtung mit dem

Kehrwert des Auswahlgesetzes der Schicht für jede Stichprobeneinheit, korrigiert um die Unit-Non-Response für jede Schicht, woraus sich die realisierte Stichprobe ergibt. Wenn der Auswahlgesetz z. B. 10/100 beträgt, sich aber 10 % der Stichprobeneinheiten nicht beteiligt haben, entspricht der korrigierte Auswahlgesetz 9/100. Dieses Verfahren beruht auf der Annahme, dass die Verteilung der Innovationsaktivitäten zwischen den antwortenden und den nicht antwortenden Einheiten identisch ist. Diese Annahme kann anhand einer Analyse der Antwortausfälle überprüft werden. Selbst wenn die Annahme falsch ist, kann die damit einhergehende Verzerrung ignoriert werden, solange die Ausfallquote (Unit-Non-Response) verhältnismäßig niedrig ist.

9.138. Nicht alle Stichprobenerhebungen werden in Schichten aufgeteilt – eine Vollerhebung ist definitionsgemäß nicht geschichtet. Im Fall einer nicht geschichteten Erhebung sollte bei der Gewichtung der Kehrwert des gesamten Auswahlgesetzes verwendet werden, bereinigt um den Kehrwert der gesamten Rücklaufquote. Bei einer Vollerhebung können im Nachhinein Schichten nach Unternehmensgröße, Sektor, Region usw. gebildet werden, um die Antwortausfallquote in der entsprechenden Schicht zu ermitteln. Die Gewichtungsvariablen für eine Vollerhebung können ausgehend vom Kehrwert der Rücklaufquote der jeweiligen Schicht erstellt werden.

9.139. Die endgültigen Gewichtungsfaktoren sollten weiter korrigiert werden, wenn eine Nicht-Teilnehmer-Befragung statistisch signifikante Unterschiede zwischen den befragten Einheiten der eigentlichen Erhebung und den befragten Einheiten der Nicht-Teilnehmer-Befragung aufzeigt. Dies ist z. B. der Fall, wenn der Prozentsatz nicht innovativer Einheiten, die an der eigentlichen Erhebung teilgenommen haben, niedriger ist als in der Nicht-Teilnehmer-Befragung. Ein Ansatz besteht darin, jede Schicht in mehrere homogene Antwortgruppen zu unterteilen, wobei die (unterstellte) Antwortwahrscheinlichkeit innerhalb jeder Gruppe gleich ist. Diese Gruppen können anhand der Ergebnisse der Nicht-Teilnehmer-Befragung festgelegt werden. Ein anderer Ansatz besteht darin, in der Schätzphase Hilfsinformationen zu verwenden, um Verzerrungen infolge von Antwortausfällen zu verringern, oder Schätzmethoden mit zweistufiger Stichprobenziehung anzuwenden. Im letzteren Fall wird die Stichprobe anhand der Merkmalsausprägung unterteilt, für die eine wahrscheinliche Antwortausfallverzerrung untersucht wurde (z. B. innovative versus nicht innovative Unternehmen). Die Gewichtungsfaktoren werden für jede Gruppe separat berechnet. In einem zweiten Schritt werden die Gewichtungsfaktoren um einen Korrekturfaktor für Antwortausfälle bereinigt, der der Verzerrung zwischen den antwortenden und den nicht antwortenden Unternehmen in Bezug auf das untersuchte Merkmal entspricht.

9.140. Die Gewichte können durch Kalibrierung weiter verfeinert werden, wenn die Auswahlgesamtheit quantitative oder qualitative Daten zu allen Einheiten enthält, wie Beschäftigtenzahl, Umsatz, Rechtsform oder Region. Die Kalibrierung gewährleistet, dass die gewichtete Stichprobe der gesamten Grundgesamtheit oder der Verteilung der Variablen in der Auswahlgesamtheit entspricht. Sie kann die Genauigkeit erhöhen und die Verzerrung reduzieren. Geeignete Softwareprogramme zur Kalibrierung sind z. B. CLAN von Statistics Sweden, CALMAR vom französischen Statistikamt INSEE oder G-Est von Statistics Canada und stehen auch anderen Ländern zur Nutzung zur Verfügung. Mit vielen der zur Berechnung der Gewichte verwendeten Softwarepakete lassen sich auch Kennzahlen zur Stichprobenvariabilität ermitteln.

9.8. Veröffentlichung und Verbreitung der Ergebnisse

9.141. Innovationserhebungen werden zur Erstellung von Tabellen mit Innovationsstatistiken und -indikatoren sowie für ökonometrische Analysen zu einer Vielzahl von Innovationsthemen verwendet. Für die Erstellung von Statistiken und Indikatoren ist eine Gewichtung der Einheiten erforderlich, um repräsentative Ergebnisse für die Grundgesamtheit zu erhalten. Die meisten Innovationserhebungen verwenden für viele Schichten eine Wahrscheinlichkeitsstichprobe. Stichprobenerhebungen können bei den Indikatoren zwei Arten von Fehlern verursachen:

Zufällige Fehler aufgrund des Zufallsverfahrens, das für die Auswahl der Einheiten angewendet wird, und systematische Fehler, die alle nicht zufälligen Fehler (Bias) umfassen. Die Wahrscheinlichkeit von Zufallsfehlern sollte zusammen mit den Ergebnissen unter Einbeziehung der Konfidenzintervalle, Standardfehler und Variationskoeffizienten angegeben werden. Die Konfidenzgrenzen geben die tatsächlichen, aber unbekanntenen Werte der Auswahlgesamtheit mit einer gegebenen Wahrscheinlichkeit wieder. Datenqualitätsberichte sollten nach Möglichkeit auch eine Evaluierung der nicht zufälligen Fehler enthalten.

9.8.1. Metadaten und Qualitätsberichte

9.142. Statistiken und Indikatoren sollten Metadaten enthalten, darunter Informationen über die Datenerhebungsmethode, Stichprobenverfahren, Verfahren zur Behandlung von Antwortausfällen und Qualitätsindikatoren. Diese Informationen ermöglichen es den Nutzern, die Daten besser zu interpretieren und ihre Qualität zu beurteilen. Internationale Organisationen sollten weiterhin detaillierte Informationen über die allgemeinen und spezifischen Methoden bereitstellen, die von den in ihren Datenbanken und Berichten erfassten Ländern angewendet werden.

9.8.2. Zugang zu den Daten

9.143. Deskriptive Daten können durch Pressemitteilungen, Tabellen, Datenbanken und Berichte bereitgestellt werden. Die ökonomische Analyse von Daten aus Innovationserhebungen ist für die Politikgestaltung von erheblichem Wert (vgl. Abschnitt 11.5), ist aber nicht die Hauptaufgabe der nationalen Statistikämter. Interne ökonomische Analysen durch nationale Statistikämter können zu geringen Kosten ausgeweitet werden, indem Forschende Zugang zu Mikrodaten aus Innovationserhebungen erhalten. Dabei müssen die Bestimmungen des Datenschutzes beachtet werden, z. B. indem entweder der Datenzugang für externe Wissenschaftler*innen über ein Forschungsdatenzentrum mit entsprechenden Datensicherheitsvorkehrungen erfolgt oder anonymisierte Datensätze bereitgestellt werden.

Literaturverzeichnis

- Barge, S. und H. Gehlbach (2012), „Using the Theory of Satisficing to Evaluate the Quality of Survey Data“, *Research in Higher Education*, Vol. 53/2, s. 182-200, <http://dx.doi.org/10.1007/s11162-011-9251-2>.
- Belfo, F. P. und R. D. Sousa (2011), „A web survey implementation framework: evidence-based design practices“, Paper für die 6. Mediterranean Conference on Information Systems, MCIS 2011, Limassol, 3.-5. September, <http://aisel.aisnet.org/mcis2011/43/>.
- Cirera, X. und S. Muzi (2016), „Measuring Firm-Level Innovation Using Short Questionnaires: Evidence from an Experiment“, *Policy Research Working Papers*, No. 7696, Weltbankgruppe, <http://documents.worldbank.org/curated/en/877171467989541697/Measuring-firm-level-innovation-using-short-questionnaires-evidence-from-an-experiment>.
- Couper, M. P. et al. (2013), „The Design of Grids in Web Surveys“, *Social Science Computer Review*, Vol. 31/3, S. 322-345, <https://doi.org/10.1177/0894439312469865>.
- Downes-Le Guin, T. et al. (2012), „Myths and realities of respondent engagement in online surveys“, *International Journal of Market Research*, Vol. 54/5, S. 613-633, <https://doi.org/10.2501/IJMR-54-5-613-633>.
- Dykema, J. et al. (2013), „Effects of E-Mailed Versus Mailed Invitations and Incentives on Response Rates, Data Quality, and Costs in a Web Survey of University Faculty“, *Social Science Computer Review*, Vol. 31/3, S. 359-370,

- <https://doi.org/10.1177/0894439312465254>.
- Europäische Kommission et al. (2009), *System of National Accounts 2008*, New York, <https://unstats.un.org/unsd/nationalaccount/docs/sna2008.pdf>.
- Fan, W. und Z. Yan (2010), „Factors affecting response rates of a web survey: A systematic review”, *Computers in Human Behavior*, Vol. 26/2, S. 132-139, <https://doi.org/10.1016/j.chb.2009.10.015>.
- Galesic, M. und M. Bosnjak (2009), „Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey”, *Public Opinion Quarterly*, Vol. 73/2, S. 349-360, <https://doi.org/10.1093/poq/nfp031>.
- Galindo-Rueda, F. und A. Van Cruysen (2016), „Testing innovation survey concepts, definitions and questions: Findings from cognitive interviews with business managers”, *OECD Science, Technology and Innovation Technical Papers*, OECD, Paris, <http://oe.cd/innocognitive>.
- Harkness, J. A. et al. (Hrsg.) (2010), *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*, Wiley Series in Survey Methodology, John Wiley & Sons, Hoboken, <http://doi.org/10.1002/9780470609927>.
- Harris, R. I. D. (1988), „Technological change and regional development in the UK: Evidence from the SPRU database on innovations”, *Regional Studies*, Vol. 22/5, S. 361-374, <https://doi.org/10.1080/00343408812331345050>.
- Hoskens, M. et al. (2016), „State of the art insights in capturing, measuring and reporting firm-level innovation indicators”, Paper für das OECD Blue Sky Forum 2016, Gent, 19.-21. September, www.oecd.org/sti/069%20-%20Measuring%20innovation_ECOOM%20August%202016.pdf.
- Kleinknecht, A., J. O. N. Reijnen und W. Smits (1993), „Collecting literature-based innovation output indicators: The experience in the Netherlands”, in A. Kleinknecht und D. Bain (Hrsg.), *New Concepts in Innovation Output Measurement*, Palgrave Macmillan, London, S. 42-84, <https://doi.org/10.1007/978-1-349-22892-8>.
- Millar, M. M. und D. A. Dillman (2011), „Improving Response to Web and Mixed-Mode Surveys”, *Public Opinion Quarterly*, Vol. 75/2, S. 249-269, <https://doi.org/10.1093/poq/nfr003>.
- OECD (2015a), *Recommendation of the OECD Council on Good Statistical Practice*, OECD, Paris, www.oecd.org/statistics/good-practice-toolkit/Brochure-Good-Stat-Practices.pdf.
- OECD (2015b), *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris, <https://doi.org/10.1787/9789264239012-en>; dt. Fassung: OECD (2018), *Frascati-Handbuch 2015: Leitlinien für die Erhebung und Meldung von Daten über Forschung und experimentelle Entwicklung*, Messung von wissenschaftlichen, technologischen und Innovationstätigkeiten, OECD Publishing, Paris, <https://doi.org/10.1787/9789264291638-de>.
- Olson, K. und B. Parkhurst (2013), „Collecting Paradata for Measurement Error Evaluations”, in *Improving Surveys with Paradata: Analytic Uses of Process Information*, John Wiley & Sons, Hoboken, S. 43-72, <https://doi.org/10.1002/9781118596869.ch3>.
- Revilla, M. und C. Ochoa (2015), „What are the Links in a Web Survey Among Response Time, Quality and Auto-Evaluation of the Efforts Done?”, *Social Science Computer Review*, Vol. 33/1, S. 97-114, <https://doi.org/10.1177/0894439314531214>.

- Saunders, M. N. K. (2012), „Web versus Mail: The Influence of Survey Distribution Mode on Employees' Response", *Field Methods*, Vol. 24/1, S. 56-73, <https://doi.org/10.1177/1525822X11419104>.
- Snijkers, G. et al. (Hrsg.) (2013), *Designing and Conducting Business Surveys*, Wiley Series in Survey Methodology, John Wiley & Sons, Hoboken, <https://doi.org/10.1002/9781118447895>.
- Snijkers, G. und D. K. Willimack (2011), „The Missing Link: From Concepts to Questions in Economic Surveys", Paper für den 2. European Establishment Statistics Workshop (EESW11), Neuenburg, Schweiz, 12.-14. September.
- Tourangeau, R., L. J. Rips und K. Rasinski (2000), *The Psychology of Survey Response*, Cambridge University Press, Cambridge, <https://doi.org/10.1017/CBO9780511819322>.
- VN (2008), *International Standard Industrial Classification of All Economic Activities (ISIC), Rev.4*, Vereinte Nationen, New York, https://unstats.un.org/unsd/classifications/Econ/Download/In%20Text/ISIC_Rev_4_publication_English.pdf.
- VN (2007), *Statistical Units*, Vereinte Nationen, New York, <http://unstats.un.org/unsd/isdts/docs/StatisticalUnits.pdf>.
- Wilhelmsen, L. (2012), „A question of context: Assessing the impact of a separate innovation survey and of response rate on the measurement of innovation activity in Norway", *Documents*, No. 51/2012, Statistics Norway, Oslo, www.ssb.no/a/english/publikasjoner/pdf/doc_201251_en/doc_201251_en.pdf.
- Willeboordse, A. (Hrsg.) (1997), *Handbook on Design and Implementation of Business Surveys*, Eurostat, Luxembourg, <http://ec.europa.eu/eurostat/ramon/statmanuals/files/Handbook%20on%20surveys.pdf>.
- Willis, G. B. (2015), *Analysis of the Cognitive Interview in Questionnaire Design*, Oxford University Press, Oxford.
- Willis, G. B. (2005), *Cognitive Interviewing: A Tool for Improving Questionnaire Design*, SAGE Publications, Thousand Oaks, CA.
- Zhang, X. C. et al. (2017), „Survey method matters: Online/offline questionnaires and face-to-face or telephone interviews differ", *Computers in Human Behavior*, Vol. 71, S. 172-180, <https://doi.org/10.1016/j.chb.2017.02.006>.



From:
Oslo Manual 2018
Guidelines for Collecting, Reporting and Using Data on
Innovation, 4th Edition

Access the complete publication at:
<https://doi.org/10.1787/9789264304604-en>

Please cite this chapter as:

OECD/Eurostat (2024), "Methoden zur Erhebung von Daten über Innovationen im Unternehmenssektor", in *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition*, OECD Publishing, Paris/Eurostat, Luxembourg.

DOI: <https://doi.org/10.1787/a4a5e67b-de>

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.