

**Handreichung AFiD-Modul CITID: Wichtige Hinweise
für Forschungsprojekte die mit den verknüpfbaren
Daten der Forschungsdatenzentren der Statistischen
Ämter des Bundes und der Länder und des
Forschungsdaten- und Servicezentrums der
Bundesbank arbeiten**

Version 1

Impressum

Herausgeber: Statistische Ämter des Bundes und der Länder
Herstellung: Information und Technik Nordrhein-Westfalen
Telefon 0211 9449-01 • Telefax 0211 9449-8000
Internet: www.forschungsdatenzentrum.de
E-Mail: forschungsdatenzentrum@it.nrw.de

Fachliche Informationen zu dieser Veröffentlichung:

Forschungsdatenzentrum des
Statistischen Bundesamts

Tel.: 0611 75-2420
Fax: 0611 75-3915
forschungsdatenzentrum@destatis.de

Informationen zum Datenangebot:

Statistisches Bundesamt
Forschungsdatenzentrum

Tel.: 0611 75-2420
Fax: 0611 75-3915

Forschungsdatenzentrum der
Statistischen Ämter der Länder
– Geschäftsstelle –

Tel.: 0211 9449-2873
Fax: 0211 9449-8087
forschungsdatenzentrum@it.nrw.de

Erscheinungsfolge: unregelmäßig
Erschienen im Juni 2024
Zuletzt aktualisiert: 10.06.2023

Diese Publikation wird kostenlos als PDF-Datei zum Download unter www.forschungsdatenzentrum.de angeboten.

© Information und Technik Nordrhein-Westfalen, Düsseldorf, 2024
(im Auftrag der Herausbergemeinschaft)

Vervielfältigung und Verbreitung, nur auszugsweise, mit Quellenangabe gestattet. Alle übrigen Rechte bleiben vorbehalten.

Fotorechte Umschlag: ©artSILENCEcom – Fotolia.com

Empfohlene Zitierung:

Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder: Handreichung AFiD-Modul CITID: Wichtige Hinweise für Forschungsprojekte die mit den verknüpfbaren Daten der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder und des Forschungsdaten- und Servicezentrums der Bundesbank arbeiten. Version 1. Wiesbaden 2024.

Inhalt

1. Allgemeine Informationen zu den Verknüpfbaren Datenprodukten	1
2. Verknüpfung des AFiD-Panel Außenhandelsstatistik (AHS)	2
a. Verknüpfung von AFiD-Panel AHS mit der Statistik zum internationalen Dienstleistungshandel (SITS) und Statistik zum internationalen Kapitalverkehr (SIFCT)	2
b. Verknüpfung von AFiD-Panel AHS mit der Mikrodatenbank Direktinvestitionen (MiDi)	7
3. Verknüpfungen mit dem AFiD-Panel Unternehmensstrukturstatistiken (SBS)	9
a. Verknüpfung des AFiD-Panel SBS mit der MiDi	9
4. Besonderheiten Ländercodes	11

1. Allgemeine Informationen zu den Verknüpfbaren Datenprodukten

In diesem Dokument sollen die wichtigsten konzeptionellen Aspekte der möglichen Datenverknüpfung und Besonderheiten der einzelnen Datenprodukte der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (FDZ) und des Forschungsdaten- und Servicezentrums (FDSZ) der Bundesbank diskutiert und empirische Eckwerte vorgestellt werden.

Alle zur Verfügung stehende Datensätze im FDZ und FDSZ liegen auf Ebene der rechtlichen Einheit vor. Wenn im Folgenden von Unternehmen die Rede ist, ist immer die rechtliche Einheit gemeint, obwohl dies nicht der Definition des Unternehmens gemäß EU-Einheiten-Verordnung (Verordnung (EWG) Nr. 696/93) entspricht.

Bei der Verknüpfung der Datensätze ist immer darauf zu achten, welcher Datensatz die Basis bildet und somit den Kerndatensatz darstellt. Abhängig davon ergibt sich welche Verknüpfungsmethode sich als sinnvoll erweist und welche Besonderheiten beachtet werden müssen.

Die Grundlage der Verknüpfung bildet das Unternehmensregister (URS) und deshalb kommt diesem bei der Verknüpfung eine Sonderrolle zu, da die Verknüpfung der Datensätze über die Kennnummer der rechtlichen Einheit aus dem URS erfolgt. Genauer gesagt kann die Verknüpfung sowohl über die Kennnummer der rechtlichen Einheit (*urs_we_id_z*), als auch die alternative Kennnummer der rechtlichen Einheit (*unr*) erfolgen. Da im URS die gleiche alternative Kennnummer der rechtlichen Einheit (*unr*) pro Jahr mehrfach vorkommen kann, ist in diesem Fall sicherzustellen, dass keine ungewollte Duplizierung der Beobachtungen erfolgt. In den meisten FDZ-Produkten der Amtlichen Firmendaten in Deutschland (AFiD-Produkte) ist nur die Kennnummer *unr* enthalten¹. Eine Anspielung der Kennnummer *urs_we_id_z* über das URS an die AFiD-Produkte, die diese nicht enthält, führt zu einer gewissen Unschärfe, da sich die *urs_we_id_z* durch Diskontinuität über die Zeit verändern kann.

Die Kennnummer der Bundesbank für die Unternehmen lautet *awmus* und ist über eine Mappingtabelle der entsprechenden Kennnummer aus dem URS zuordenbar. Die Verknüpfung erfolgt über eine Verknüpfungsdatei, die als AFiD-Modul CITID (Combined International Trade and Investment Data) im FDZ verfügbar ist. Mehr Informationen und wichtige Hinweis beim Zusammenführen der Kennnummern finden Sie im Metadatenreport Teil Produkt des *AFiD-Modul Combined International Trade and Investment Data*². Die Verknüpfbarkeit zweier Datensätze ist unmittelbar davon abhängig, wie gut die Datensätze

¹ Nähere Informationen zu den enthaltenen Kennnummern und Variablen in den einzelnen AFiD-Datenprodukten sind auf der Homepage der Forschungsdatenzentren des Bundes und der Länder unter „weitere Informationen“ der einzelnen Datenprodukte zu finden (siehe <https://www.forschungsdatenzentrum.de/de/afid>).

² LINK einfügen

sich jeweils mit dem URS verknüpfen lassen. Aus diesem Grund wird die Verknüpfung der einzelnen Datensätze in diesem Dokument zur Hilfestellung beleuchtet. Dabei wird speziell auf die Verknüpfung der FDZ-Produkte AFiD-Panel Außenhandelsstatistik (AHS) und AFiD-Panel Unternehmensstrukturstatistiken (Structural Business Statistics -SBS) mit den FDSZ-Produkten Mikrodatenbank Direktinvestitionen (MiDi), Statistik zum internationalen Dienstleistungshandel (SITS) und Statistik zum internationalen Kapitalverkehr (SIFCT) eingegangen.

2. Verknüpfung des AFiD-Panel Außenhandelsstatistik (AHS)

a. Verknüpfung von AFiD-Panel AHS mit der Statistik zum internationalen Dienstleistungshandel (SITS) und Statistik zum internationalen Kapitalverkehr (SIFCT)

Die Datensätze AFiD-Panel AHS, SITS und SIFCT enthalten im weiteren Sinne Transaktionsdaten, die jeweils auf unterschiedlicher Ebene disaggregiert werden. In allen drei Datensätzen werden Transaktionen nach den folgenden Merkmalen differenziert:

- **Unternehmen:** Der Datensatz AFiD-Panel AHS enthält die Kennnummer der rechtlichen Einheit aus dem URS. Dieselbe Kennnummer lässt sich an SIFCT und SITS durch die Mappingtabelle (AFiD-Modul CITID) der Bundesbank basierend auf dem URS 2014 bis 2021 anspielen.
- **Verkehrsrichtung:** In den Datensätzen SIFCT und SITS gibt es für die beiden Verkehrsrichtungen jeweils eine Spalte. Im AFiD-Panel AHS hingegen gibt es für jede Verkehrsrichtung eine eigene Beobachtung. Dies muss bei einer Verknüpfung vereinheitlicht werden. Darüber hinaus stellt sich die Frage, ob mit Blick auf die Verkehrsrichtung der Fokus auf den Zahlungsströmen liegt oder auf den Im- bzw. Exporten liegt (siehe unten Zahlungsströme vs. Warenströme). Grundsätzlich gilt, dass Waren-, Dienstleistungs-, und Kapitalimporte als Ausgaben zu werten sind (oder umgekehrt) und Exporte als Einnahmen (oder umgekehrt).
- **Partnerland:** Alle drei Datensätze enthalten Informationen zum Partnerland als zweistellige alphanumerische ISO-Codes. Zu beachten sind die Ausführungen zum Unterschied zwischen Zahlungsströmen und Warenströmen. Darüber hinaus ist zu beachten, dass die Außenhandelsstatistik bei Importen sowohl das Versendungsland (Land aus dem die Ware versendet wurde), als auch das Ursprungsland (Land in dem die letzte nennenswerte Änderung an der Ware vorgenommen wurde) ausweist. Hier obliegt es den Nutzenden, sich für das Versendungsland- oder Ursprungslandprinzip

zu entscheiden³. Die nationalen Veröffentlichungen im Außenhandel folgen dem Ursprungslandprinzip. Darüber hinaus gibt es Besonderheiten und Unterschiede bei den Ländercodes, die in Anhang A näher beschrieben sind.

- **Klassifikationen:** Die Angaben aus dem AFiD-Panel AHS werden nach dem Warenverzeichnis für die Außenhandelsstatistik⁴ differenziert, die Transaktionen aus SIFCT und SITS auf Basis der Kennzahlen aus dem Leistungsverzeichnis der Deutschen Bundesbank.
- **Monat-Jahr:** Alle drei Datensätze enthalten Informationen zum Bezugsmonat. In SIFCT und SITS ist der Bezugsmonat der Monat der Zahlung, im AFiD-Panel AHS ist der Bezugsmonat der Monat der Warenbewegung (siehe Zahlungsströme vs. Warenströme).

Grundsätzlich gibt es, darauf aufbauend, zwei Möglichkeiten die Daten sinnvoll miteinander zu verknüpfen, an denen sich Forschende orientieren können.

- I. Die **erste Option** besteht darin, die Daten in einem Datensatz im Long-Format zusammenzuführen. In diesem Fall würden Beobachtungen aus allen drei Datensätzen untereinander in einem Datensatz zusammengeführt und um eine Angabe zum Quelldatensatz ergänzt. Dabei müssen die drei Datensätze in einem ersten Schritt so harmonisiert werden, dass die oben beschriebenen Variablen sinnvoll zusammenpassen. Was dabei zu beachten ist, wird unten ausgeführt⁵. Darüber hinaus sind Entscheidungen erforderlich, wie mit Merkmalen umzugehen ist, die nicht in allen Datensätzen vorkommen. Es kann sinnvoll sein, die einzelnen Datensätze zunächst so zu aggregieren – also die Wertangaben entsprechend zu summieren –, dass je Kombination der genannten Variablen in jedem der Quelldatensätze nur eine Beobachtung übrigbleibt.
- II. Die **zweite Option** besteht darin, die Daten über Identifikatoren zu verknüpfen. Auch hier sollten die einzelnen Datensätze zunächst so aggregiert werden, dass keine Duplikate mehr in den jeweils gewünschten Verknüpfungsvariablen bestehen, da ansonsten ein kartesisches Produkt der Datensätze erzeugt wird und potentiell Beobachtungen aus allen Datensätzen dupliziert würden. In Einzelfällen kann es sinnvoll sein, Duplikate in einem einzigen Datensatz zu belassen, wenn der dadurch entstehenden Duplizierung von Werten in den anderen Datensätzen im Folgenden

³ Siehe Abschnitt Partnerland:

https://www.destatis.de/DE/Themen/Wirtschaft/Aussenhandel/Methoden/Downloads/infoblatt-asymmetrien.pdf?__blob=publicationFile.

⁴ Siehe Link: <https://www.destatis.de/DE/Methoden/Klassifikationen/Aussenhandel/warenverzeichnis-aussenhandel-2023.html>.

⁵ Siehe Abschnitte: Zahlungsströme vs. Warenströme, Abgrenzung Güter und Dienstleistungen, Periodizität und Datentiefe und Repräsentativität (Seite 4-6)

Rechnung getragen wird. Die Verknüpfung ist grundsätzlich auf verschiedenen Ebenen und über verschiedene Kombinationen der folgenden Merkmale möglich:

- **Unternehmen:** Ein Match ist über die Kennnummer der rechtlichen Einheit aus dem URS möglich.
- **Verkehrsrichtung:** Werden die einzelnen Datensätze in einem ersten Schritt so harmonisiert, dass alle Datensätze eine Variable zur Verkehrsrichtung enthalten, so ist eine Verknüpfung über diese Variable unter Berücksichtigung der oben genannten Besonderheiten zur jeweiligen Bedeutung möglich.
- **Partnerland:** Alle drei Datensätze enthalten Informationen zum Partnerland als zweistellige alphanumerische ISO-Codes, beachte die Abweichungen in Kapitel 4. Auch über diese Variable ist eine Verknüpfung möglich, nachdem sich die Nutzenden für das Versendungsland- oder Ursprungslandprinzip in den Außenhandelsdaten entschieden haben.
- **Monat-Jahr:** Alle drei Datensätze enthalten Informationen zum Bezugsmonat. In SIFCT und SITS ist der Bezugsmonat der Monat der Zahlung, im AFiD-Panel AHS ist der Bezugsmonat der Monat der Warenbewegung. (Siehe Zahlungsströme vs. Warenströme).

Über die Klassifikationsvariablen (Warennummer aus dem Warenverzeichnis für die Außenhandelsstatistik und Kennzahl aus dem Leistungsverzeichnis der Deutschen Bundesbank) ist keine Verknüpfung möglich, da jeweils unterschiedliche Dinge (Waren, Dienstleistungen, Kapitalverkehre) klassifiziert werden.

Bei beiden Verknüpfungsoptionen ist zusätzlich folgendes zu beachten:

- **Zahlungsströme vs. Warenströme:** Es ist zwischen Zahlungs- und Warenströmen zu unterscheiden. Grundsätzlich enthalten die SITS und SIFCT Daten der Bundesbank Informationen zu grenzüberschreitenden Zahlungsströmen. Die Daten der Außenhandelsstatistik hingegen bilden den Wert der Waren bei physischem Grenzüberschritte ab. Dies kann einem Zahlungsstrom entsprechen. In der Regel weicht der Wert der Ware aber vom tatsächlichen Zahlungsstrom ab. Ein wichtiger Grund dafür sind die Lieferbedingungen: Die Zahlen des AFiD-Panels AHS beinhalten Transportkosten bis zur deutschen Grenze unabhängig davon, wer diese bezahlt und wer die Transportleistung zur Verfügung stellt. Je nach Lieferbedingungen kann der Zahlungsstrom demnach höher oder niedriger ausfallen als der Wert der Ware an der deutschen Grenze. Ein weiterer Grund für Abweichungen ist, dass bei manchen Geschäftsarten keine Zahlung in Höhe des Wertes der Ware erfolgt. Das ist zum Beispiel dann der Fall, wenn die Ware zwar bewegt wird, aber nicht den Eigentümer

wechselt oder wenn eine Ware unentgeltlich versendet wird. Der Wert der Ware in der Außenhandelsstatistik wäre also größer als der tatsächliche Zahlungsstrom. Zusätzlich ist zu beachten, dass der Zahlungsstrom, der im Rahmen einer Warenlieferung erfolgt, in der Regel in die entgegengesetzte Richtung fließt. Sofern eine Zahlung getätigt werden muss, entspricht einem Wareneingang (Import) ein Zahlungsausgang (Ausgabe) und umgekehrt. In diesem Zusammenhang ist auch zu beachten, dass im Rahmen von Dreiecksgeschäften das Partnerland aus dem AFiD-Panel AHS nicht dem Land entspricht, in das die Zahlung fließt. Wenn z.B. ein französisches Unternehmen eine Ware an ein deutsches Unternehmen verkauft und diese direkt aus Belgien liefern lässt, so entsteht eine Warenbewegung aus Belgien nach Deutschland. Dass die entsprechende Zahlung von Deutschland nach Frankreich fließt, ist aus den Daten des AFiD-Panels AHS nicht ersichtlich. Auch zeitlich kann die Zahlung vom Warenstrom abweichen, ohne dass dies aus den Daten hervorgeht.

- **Abgrenzung Güter und Dienstleistungen:** Es gibt Transaktionen, die sowohl in SITS als auch in der Außenhandelsstatistik erfasst werden, wobei die Behandlung jedoch jeweils abweicht. Dies ist zum einen bei Lohnveredelung der Fall. Lohnveredelung bezeichnet die Be- oder Verarbeitung von Waren, die sich nicht im Eigentum des bearbeitenden Unternehmens befinden. Grundsätzlich handelt es sich dabei um eine Dienstleistung. Sofern Auftraggeber und Auftragnehmer ihren Sitz in unterschiedlichen Ländern haben, handelt es sich um eine internationale Dienstleistung, die in Höhe des Lohnveredelungsentgelts in der SITS erfasst wird. Gleichzeitig kann es in diesem Zusammenhang zu Warenbewegungen kommen, die im AFiD-Panel AHS enthalten sind, wenn unentgeltlich Vormaterialien geliefert werden oder das fertige Produkt das Land wieder verlässt. Auch der im AFiD-Panel AHS enthaltene Wert des fertigen Produkts enthält im letzteren Fall das Lohnveredelungsentgelt. Um zu vermeiden, dass das Lohnveredelungsentgelt doppelt gezahlt wird, sollten Lohnveredelungen nur aus einer der beiden Quellen berücksichtigt werden.⁶ Darüber hinaus variiert die Zuordnung. In SITS sind die Lohnveredelungsgeschäfte inländischen Unternehmen zugeordnet. Gegenüber der Außenhandelsstatistik sind hingegen bei Lohnveredelung im Inland teilweise die ausländischen Auftraggeber meldepflichtig. Diese Transaktionen sind im AHS-Panel nicht enthalten. Neben Lohnveredelung ergibt sich für Transportdienstleistungen eine weitere Besonderheit. Wie unter Punkt

⁶ Zu beachten ist allerdings, dass Lohnveredelung im Sinne der SITS auch ohne jegliche grenzüberschreitende Warenbewegung im AHS erfolgen kann, zum Beispiel, wenn die Dienstleistenden an den Ort der Ware reisen und die Arbeit dort vollständig verrichtet wird. Darüber hinaus könnte zum Beispiel ein deutscher Auftraggeber im Ausland Vormaterialien erwerben, diese dort veredeln lassen und anschließend das fertige Produkt in ein anderes Land weiterverkaufen, ohne dass jemals die deutsche Grenze überschritten würde. Das Lohnveredelungsentgelt würde dann nur in der SITS erfasst.

Zahlungsströme vs. Warenströme angesprochen, enthalten die in der Außenhandelsstatistik erfassten Werte Transportleistungen bis zur deutschen Außengrenze.

- **Periodizität und Datentiefe:** Die SITS- und SIFCT-Daten der Deutschen Bundesbank sind wie die Daten des AFiD-Panels AHS nach Partnerländern, Monat und Verkehrsrichtung differenziert. Alle drei Datensätzen enthalten zusätzliche Klassifikationsvariablen. Die Unterscheidung zwischen Versendungs- und Ursprungsland, sowie Angaben zu Verkehrszweig, Präferenznachweis, Art des Geschäfts und inländische Ursprungs-/Zielland gibt es nur im AFiD-Panel AHS.
- **Repräsentativität:** Es handelt sich sowohl bei der Außenhandelsstatistik, als auch bei allen Datensätzen der Bundesbank um Vollerhebungen mit Abschneidegrenze. Die Abschneidegrenzen sind jeweils spezifisch für die einzelne Statistik und beziehen sich jeweils auf unterschiedliche Größen. In SITS und SIFCT bezieht sich die Meldepflicht auf die Transaktion und somit auf die Kombination von Unternehmen, Monat, Partnerland und Kennziffer.⁷ Folglich können für dasselbe Unternehmen im selben Zeitraum einige Arten von Transaktionen meldepflichtig sein und andere nicht. In der Außenhandelsstatistik hingegen bezieht sich die Meldeschwelle auf Unternehmen und Jahr. Des Weiteren werden in SITS einige Dienstleistungen nicht erhoben, sondern geschätzt, wie beispielsweise der Reiseverkehr und Ausgaben für den digitalen Handel.

Darüber hinaus ist bei der ersten Verknüpfungsoption eines Datensatzes im Long-Format Folgendes zu beachten:

- **Differenzierung nach Transaktionen:** In keiner der Quellen mit Flussgrößen liegen echte Transaktionsdaten vor, d.h. Daten zu einzelnen Rechnungen oder Geschäften von Unternehmen. Die Beobachtungen sind lediglich nach den jeweils vorhandenen kategorialen Merkmalen differenziert. Es können also keine Rückschlüsse auf die Zahl der Transaktionen gezogen werden. In diesem Zusammenhang sind die Ausführungen unten zu den verwendeten Klassifikationen zu beachten.
- **Klassifikationen:** In den AFiD-Panel AHS Daten, den SITS-Daten und den SIFCT-Daten lassen sich verschiedene Arten von Transaktionen unterscheiden. Das AFiD Panel AHS differenziert Waren nach dem Warenverzeichnis für die

⁷ Siehe Biewen, E., L. Pham-Dao and H. Stahl (2022). Statistics on international financial and capital transactions (SIFCT), Data Report 2022-07 – Metadata Version 4. Deutsche Bundesbank, Research Data and Service Centre, Seite 5. <https://www.bundesbank.de/resource/blob/889656/f527fc74b221787a1bc958413cb0ecc7/mL/2022-07-sifct-data.pdf>

Außenhandelsstatistik⁸. In den SITS- und SIFCT-Daten werden Transaktionen nach den Kennzahlen aus dem Leistungsverzeichnis der Deutschen Bundesbank für die Zahlungsbilanzstatistik⁹ differenziert. Die verwendeten Klassifikationen unterscheiden sich teils stark in Ihrer Detailtiefe. Das Warenverzeichnis für die Außenhandelsstatistik enthielt 2020 insgesamt 9.483 achtstellige Warennummern, das Leistungsverzeichnis der Deutschen Bundesbank enthält 103 dreistellige Kennziffern für Dienstleistungen und 173 dreistellige Kennziffern für Kapitalverkehre und -erträge. Mit Blick auf SITS und AFiD-Panel AHS existieren grundsätzlich Klassifikationen, die sowohl Güter als auch Dienstleistungen umfassen und damit grundsätzlich eine Vereinheitlichung ermöglichen.¹⁰ Eine Überführung der SITS-Daten in diese Klassifikationen ist jedoch ohne starke Annahmen nicht möglich. Die genannten Unterschiede zwischen den Klassifikationen erschweren eine Analyse auf Detailebene. Darüber hinaus ergeben sich Änderungen über die Zeit. Das Warenverzeichnis für die Außenhandelsstatistik wird jedes Jahr aktualisiert und geändert. Das Leistungsverzeichnis der Deutschen Bundesbank richtet sich nach dem Balance of Payment Manual und wurde zuletzt 2014 grundlegend angepasst.

b. Verknüpfung von AFiD-Panel AHS mit der Mikrodatenbank Direktinvestitionen (MiDi)

Der Datensatz MiDi lässt sich mit dem AFiD-Panel AHS (und evtl. angefügten SITS und SIFCT Daten) über mehrere Dimensionen verknüpfen. Dabei ist zu berücksichtigen, dass je nach Auswahl der Verknüpfungsvariablen in beiden Datensätzen mehrere Beobachtungen je Kombination der Verknüpfungsvariablen enthalten sein können. Um zu vermeiden, dass dies bei einer Verknüpfung zu einer Vervielfältigung der Beobachtungen in den Datensätzen führt, sollte mindestens einer der beiden Datensätze vorab aggregiert werden.

Die Verknüpfung ist grundsätzlich über verschiedene Kombinationen der folgenden Merkmale möglich:

- **Unternehmen:** Der MiDi Datensatz stellt Beziehungen zwischen Unternehmen dar, enthält also grundsätzlich IDs für mehrere Unternehmen je Beobachtung. Die Variable *num* enthält bei deutschen Direktinvestitionen im Ausland die ID der deutschen Einheit, die indirekt oder direkt die entsprechenden Aktiva hält. Bei ausländischen

⁸ Siehe <https://www.destatis.de/DE/Methoden/Klassifikationen/Aussenhandel/warenverzeichnis-aussenhandel-2023.html>

⁹ Siehe <https://www.bundesbank.de/de/service/meldewesen/leistungsverzeichnis-der-deutschen-bundesbank-fuerdie-zahlungsbilanzstatistik-anlage-lv-zur-aussenwirtschaftsverordnung--611566>. Das Leistungsverzeichnis beruht seinerseits auf der Extended Balance of Payments Service Classification (EBOPS).

¹⁰ Z.B. die Classification of Products by Activity (CPA) und die Klassifikation der Broad Economic Categories Rev. 5 (BEC).

Direktinvestitionen im Inland enthält *num* die ID der deutschen Einheit, an der sich ausländische Unternehmen beteiligt haben. Grundsätzlich stehen IDs für Unternehmen in weiteren Rollen zur Verfügung. In allen Fällen können für die deutschen Einheiten IDs der rechtlichen Einheit aus dem Unternehmensregister zugespielt werden. Der Datensatz AFiD-Panel AHS enthält die ID der rechtlichen Einheit aus dem URS bereits.

- **Verkehrsrichtung:** Im MiDi werden sowohl ausländische Beteiligungen an deutschen Unternehmen, als auch deutsche Beteiligungen an ausländischen Unternehmen erfasst. Grundsätzlich kann man hier von Verkehrsrichtungen sprechen (Inward FDI vs. Outward FDI). Während sich darüber grundsätzlich eine Verknüpfung mit den AFiD-Panel AHS Daten erreichen lässt, gestaltet sich die Interpretation womöglich als schwieriger, da Bestands- und Flussgrößen zusammengeführt werden. Ob eine Verknüpfung über die Verkehrsrichtung sinnvoll ist, sollte daher sorgfältig nach Forschungsfrage geprüft werden.
- **Partnerland:** Der MiDi Datensatz enthält Informationen zum Land aller ausländischer Unternehmen, die am jeweiligen Beziehungsgeflecht beteiligt sind. Es obliegt den Nutzenden hier die Rolle zu identifizieren, die für die Forschungsfrage die größte Relevanz besitzt. Grundsätzlich ließe sich Informationen zu Direktinvestitionen aus einem Land aus MiDi aber mit Informationen zu Im- und Exporten aus diesem Land verknüpfen. Auch hier erfolgt die Angabe gemäß zweistelligen alphanumerischen ISO-Codes. Darüber hinaus ist zu beachten, dass die Außenhandelsstatistik bei Importen sowohl das Versendungsland (Land aus dem die Ware versendet wurde), als auch das Ursprungsland (Land in dem die letzte nennenswerte Änderung an der Ware vorgenommen wurde) ausweist. Hier obliegt es den Nutzenden sich für das Versendungsland- oder Ursprungslandprinzip zu entscheiden. Die nationalen Veröffentlichungen folgen dem Ursprungslandprinzip. Darüber hinaus gibt es Besonderheiten und Unterschiede bei den Ländercodes, die in Anhang A näher beschrieben sind.
- **Jahr:** MiDi-Daten sind lediglich auf Jahresebene verfügbar. Monatliche Variation gibt es nur im AFiD-Panel AHS, SITS und SIFCT.

Des Weiteren ist bei der Verknüpfung folgendes zu beachten:

- **Repräsentativität MiDi:** Im MiDi Datensatz sind definitionsgemäß alle Unternehmensbeteiligungen deutscher Unternehmen im Ausland oder ausländischer Unternehmen an deutschen Unternehmen, die die Relevanzkriterien für ausländische Direktinvestitionen erfüllen, enthalten. Gemäß dieser Relevanzkriterien sind Unternehmensbeteiligungen nur dann zu melden, wenn das Unternehmen, welches

Gegenstand einer Direktinvestition ist, eine Bilanzsumme von mindestens 3 Mio. Euro aufweist und mehr als 10% der Anteile oder Stimmrechte vom Investor gehalten werden.

- **Repräsentativität AFiD-Panel AHS:** Es handelt sich bei der Außenhandelsstatistik um eine Vollerhebung mit Abschneidegrenze. Alle außenhandelsaktiven deutschen Unternehmen sind im Datensatz enthalten. Ausländische Firmen fehlen im AFiD-Panel AHS.
- **Periodizität und Datentiefe:** Bei MiDi-Daten handelt es sich um jährliche Daten. Die Daten aus dem AFiD-Panel AHS haben eine monatliche Periodizität. Grundsätzlich besitzt der Datensätze AFiD-Panel AHS zusätzliche kategoriale Merkmale, die Analysen auf unterschiedlicher Detailebene erlauben. Eine Verknüpfung kann zur Duplizierung der MiDi Beobachtungen führen. Dies muss bei der weiteren Bearbeitung und Analyse der Daten berücksichtigt werden.

3. Verknüpfungen mit dem AFiD-Panel Unternehmensstrukturstatistiken (SBS)

a. Verknüpfung des AFiD-Panel SBS mit der MiDi

Grundsätzlich lassen sich die MiDi Daten mit den Daten aus dem AFiD-Panel SBS über die Kennnummer der rechtlichen Einheit aus dem URS in verschiedener Weise auf Unternehmensebene verknüpfen, da in der MiDi Unternehmen in unterschiedlichen Rollen auftauchen:

- **Unternehmen:** Die MiDi stellt Beziehungen zwischen Unternehmen dar, enthält also grundsätzlich IDs für mehrere Unternehmen je Beobachtung. Die Variable *num* enthält bei deutschen Direktinvestitionen im Ausland die ID der deutschen Einheit, die indirekt oder direkt die entsprechenden Anteile oder Stimmrechte an der jeweiligen ausländischen Einheit hält. Bei ausländischen Direktinvestitionen im Inland enthält *num* die ID der deutschen Einheit, an der sich ausländische Unternehmen beteiligt haben. Grundsätzlich stehen IDs für Unternehmen in weiteren Rollen zur Verfügung. So kann bei indirekten Beteiligungen an deutschen Unternehmen z.B. auch das deutsche Unternehmen, durch das die Beteiligung besteht, (identifiziert in diesen Fällen durch *nu2*) ausgewertet werden oder das deutsche Unternehmen, dass die letzte Kontrolle im Rahmen einer Kette indirekter Beteiligungen innehat (identifiziert durch *nu1*). Für den Großteil der deutschen Einheiten können IDs der rechtlichen Einheit aus dem Unternehmensregister zugespielt werden. Diese ist auch im AFiD-Panel SBS enthalten und ermöglicht eine einfache Verknüpfung.

Außerdem muss bei der Verknüpfung folgendes beachtet werden:

- **Aggregation MiDi:** Die Daten aus MiDi enthalten allgemeine Informationen sowohl zu den Unternehmen, an denen Anteile gehalten werden, als auch zu den Investoren und zu den Beziehungen zwischen den Einheiten. Einige Variablen enthalten Anteile, bei denen sich der Zähler auf das Unternehmenspaar, der Nenner aber entweder auf den Investor oder das Investitionsobjekt beziehen. Dies muss bei einer Aggregation beachtet werden, unabhängig davon warum diese erfolgt. Eine Summierung der Variablen ergibt deshalb nicht immer Sinn.
- **Repräsentativität SBS:** Der SBS Datensatz umfasst mehrere unterschiedlich rotierende Stichproben. Auswertung über Merkmale aus mehreren Quellen sind nur für die Stichprobe sinnvoll. Hochrechnungen der Wertangaben aus den drei Datensätzen auf Basis der Stichprobe stimmen **nicht** mit den veröffentlichten Zahlen überein. Dies gilt im Besonderen für die Länderebene.
- **Repräsentativität MiDi:** Im MiDi Datensatz sind definitionsgemäß alle Unternehmensbeteiligungen deutscher Firmen im Ausland und ausländischer Unternehmen an deutschen Unternehmen, die die Relevanzkriterien für ausländische Direktinvestitionen erfüllen¹¹, enthalten. Gemäß dieser Relevanzkriterien sind Unternehmensbeteiligungen nur dann zu melden, wenn das Unternehmen, das Gegenstand einer Direktinvestition ist, eine Bilanzsumme von mindestens 3 Mio. Euro aufweist und mehr als 10% der Anteile oder Stimmrechte vom Investor gehalten werden.
- **Periodizität und Datentiefe:** Die Daten des AFiD-Panels SBS enthalten pro Jahr und Unternehmen nur jeweils eine Beobachtung. In den MiDi-Daten sind jedoch komplexere Beziehungen zwischen Unternehmen dargestellt. Je deutscher Einheit und Jahr kann es dabei eine Vielzahl von Beobachtungen geben.
- **Überschneidungen von Variablen:** Der MiDi Datensatz enthält Informationen zu Wirtschaftszweig, Rechtsform, zum Sitzbundesland der deutschen Einheit, dem Umsatz und der Anzahl an Beschäftigten. Diese Variablen sind auch im SBS-Panel enthalten. Aufgrund von Unterschieden im Zeitpunkt der Erfassung, unterschiedlichen Definitionen und Kriterien kann es zu Differenzen zwischen den Angaben in den Datenquellen kommen. Häufig gehen Unternehmen mehreren Tätigkeiten nach, die unterschiedlichen Wirtschaftszweigen zugeordnet werden. Die in den Datensätzen enthaltenen Wirtschaftszweige beziehen sich in der Regel auf die Hauptaktivität des

¹¹ Siehe Friederich, K., L. Pham-Dao, C.-J. Schild, D. Scholz und J. Schumacher (2021). Microdatabase Direct Investment, Data Report 2021-23 – Document Version 1-0. Deutsche Bundesbank, Research Data and Service Center, Seite 5f.
<https://www.bundesbank.de/resource/blob/745186/8d5a001f3fba76a7b561bb13e3c53821/mL/2021-23-midi-data.pdf>.

Unternehmens. So kann es beispielweise auf Grund unterschiedlicher Kriterien, wie die Hauptaktivität eines Unternehmens zu bestimmen ist, unterschiedliche Angaben zum Wirtschaftszweig eines Unternehmens in den Datensätze geben.

- **Auslandskontrollierte Unternehmen in Deutschland:** Das AFiD-Panel SBS enthält Informationen zu ausländisch kontrollierten Unternehmen in Deutschland. Die Definition von ausländischer Kontrolle weicht von der Definition von ausländischen Unternehmensbeteiligungen in MiDi ab. Im AFiD-Panel SBS wird ein in Deutschland ansässiges Unternehmen dann als ausländisch kontrolliert betrachtet, wenn die ausländische Muttergesellschaft – direkt oder indirekt – mehr als 50% der Stimmrechte innehat bzw. faktisch Kontrolle ausübt. Dies ist die Grundlage der Statistik über auslandskontrollierte Unternehmen.¹² Im MiDi Datensatz geht es um Investitionen, aber nicht unbedingt um Kontrolle. Daher sind auch Beziehungen enthalten, die diese Definition von Kontrolle nicht erfüllen. Bereits ein Anteil von 10% der Stimmrechte in ausländischer Hand genügt für eine Erfassung in MiDi.¹³ Gemäß der Relevanzkriterien für MiDi, sind Unternehmen mit einer Bilanzsumme von weniger als 3 Mio. Euro nicht enthalten. Diese können aber im Sinne des AFiD-Panels SBS ausländisch kontrolliert und als solche gekennzeichnet sein. Da die Relevanzschwellen in MiDi kleine und viele mittelständische Unternehmen ausschließen, ist die Zahl der ausländisch kontrollierten Unternehmen im Unternehmensregister höher als die Zahl der in MiDi erfassten deutschen Unternehmen mit ausländischer Beteiligung.

4. Besonderheiten Ländercodes

Bei der Verknüpfung des AFiD-Panels AHS mit den Bundesbankdaten SITS und SIFCT über die Partnerländer müssen folgende Unterschiede in den Merkmalen der Ländercodes (*ahs_bld_uld* vs. *lan*) berücksichtigt werden:

- **Unbekannte Länder:** In den Datensätzen der Bundesbank und dem AFiD-Panel AHS wird das Kürzel „QU“ für unbekannte Länder verwendet. In den Daten der Deutschen Bundesbank wird für Internationale Organisationen das Kürzel „1A“ verwendet. Dies ist im AFiD-Panel AHS nicht enthalten, stattdessen wird jeweils das Land, in dessen Gebiet die Internationale Organisation ihren Sitz hat, gemeldet. Zusätzlich gibt es in SITS und SIFCT ein Kürzel für Polargebiete („G2“), in der Außenhandelsstatistik werden diese seit 2000 nicht mehr gesondert ausgewiesen. Im AFiD-Panel AHS gibt es stattdessen folgende besondere Codes:

¹² Siehe Destatis (2022), Statistik über auslandskontrollierte Unternehmen. Qualitätsbericht. https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Unternehmen/auslandsunternehmenseinheiten.pdf?__blob=publicationFile

¹³ Siehe Friederich et al. (2021), S. 5.

- QQ: Schiffs- und Luftfahrtbedarf
- QP: Hohe See
- **Französische Überseegebiete** werden in SIFCT und SITS teils gesondert ausgewiesen. Sie werden mittels ISO 2-Steller Code „GF“ (Französisch Guyana) und „YT“ (Mayotte) identifiziert und im AFiD-Panel AHS Frankreich (Code „FR“) zugeordnet.
- **Ozeanien:** Im SITS und SIFCT werden gesonderte ISO-Codes für US-amerikanische Gebiete („F9“), australische Gebiete („F8“) und neuseeländische Gebiete („G1“) verwendet, die keine Entsprechung im AFiD-Panel AHS haben.
- **Britische Kanalinseln** werden in SIFCT und SITS gesondert ausgewiesen unter ISO 2-Steller „GG“ (Guernsey), „IM“ (Isle of Man) und „JE“ (Jersey), im AFiD-Panel AHS aber dem Vereinigten Königreich (Code „GB“) zugeordnet.
- Für die **Niederländische Karibik** gibt es einen umfassenden ISO-Code (Code „AN“) in den Daten der Deutschen Bundesbank. Dieser ist nicht teil der Klassifikation der Außenhandelsstatistik und somit auch nicht in den Daten enthalten. Zusätzlich gibt es sowohl in den Datenreports der Bundesbank, als auch in der Außenhandelsstatistik einen gesonderten ISO-Code, der die besonderen Gemeinden der Niederlande Bonaire, Sint Eustatius und Saba umfasst (Code „BQ“). Für die autonomen Länder der ehemaligen Niederländischen Antillen gibt es in allen Quellen den Code „CW“ für Curaçao, „AW“ für Aruba und „SX“ für Sint Maarten.