

ENOB: dataNWG

## Forschungsdatenbank Nichtwohngebäude

Förderkennzeichen 03 ET1315 A

**Institut Wohnen  
und Umwelt GmbH**

Forschungseinrichtung  
des Landes Hessen und  
der Stadt Darmstadt

Rheinstraße 65  
64295 Darmstadt  
Germany

Tel: +49 (0)6151 / 2904-0  
Fax: +49 (0)6151 / 2904-97

info@iwu.de  
www.iwu.de



Leibniz-Institut  
für ökologische  
Raumentwicklung



**BERGISCHE  
UNIVERSITÄT  
WUPPERTAL**

E.1.4.5

## Stichprobe: Modellierung und Ziehung

Gefördert vom

**Bundesministerium für Wirtschaft und Energie**

im Förderbereich

**Energieoptimierte Gebäude und Quartiere**

der Fördermaßnahme *Anwendungsorientierte nichtnukleare FuE* des 6. Energie-  
forschungsprogramms der Bundesregierung



Bundesministerium  
für Wirtschaft  
und Energie

6. Mai 2021

## Impressum

|                              |                                                                                                                                                                                                                                           |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Projekt                      | <b>Forschungsdatenbank Nichtwohngebäude.<br/>Repräsentative Primärdatenerhebung zur statistisch<br/>validen Erfassung und Auswertung der Struktur und<br/>der energetischen Qualität des<br/>Nichtwohngebäudebestands in Deutschland.</b> |
| Kurztitel                    | <b>ENOB:dataNWG</b>                                                                                                                                                                                                                       |
| Teilprojekt                  | Repräsentative Stichprobenerhebung und Auswertung<br>typologischer, struktureller und energetischer Merkmale                                                                                                                              |
| Gefördert mit<br>Mitteln von | Bundesministerium für Wirtschaft und Energie (BMWi)<br>im Förderschwerpunkt Energieoptimierte Gebäude und<br>Quartiere des 6. Energieforschungsprogramms der<br>Bundesregierung                                                           |
| Förderkennzeichen            | 03 ET1315 A                                                                                                                                                                                                                               |
| Fördermittelnehmer           | Institut Wohnen und Umwelt<br>Rheinstraße 65<br>64295 Darmstadt<br>Tel. +49 (0) 6151 / 2904 -0                                                                                                                                            |
| Projektpartner               | <ul style="list-style-type: none"><li>• Leibniz-Institut für ökologische Raumentwicklung<br/>(IÖR)</li><li>• Bergische Universität Wuppertal, Fachgebiet<br/>Ökonomie des Planens und Bauens (BUW-ÖPB)</li></ul>                          |
| Auftragnehmer                |                                                                                                                                                                                                                                           |
| Bericht                      | E 1.4.5                                                                                                                                                                                                                                   |
| Verfasser                    | Dr. Holger Cischinsky                                                                                                                                                                                                                     |
| Mitarbeit                    | Michael Hörner, Markus Rodenfels                                                                                                                                                                                                          |
| Datum                        | 6. Mai 2021                                                                                                                                                                                                                               |

## Inhalt

|       |                                                                                                    |    |
|-------|----------------------------------------------------------------------------------------------------|----|
| 1     | Zum Aufbau des Berichts .....                                                                      | 1  |
| 2     | Anforderungen an das stichprobentheoretische Modell .....                                          | 3  |
| 3     | Stichprobendesign und Schätzung .....                                                              | 6  |
| 3.1   | Auswahlgrundlage .....                                                                             | 6  |
| 3.1.1 | Defizit der Auswahlgrundlage: Fehlende Informationen zu<br>Gebäudeansprechpartnern .....           | 9  |
| 3.1.2 | Defizit der Auswahlgrundlage: Fehlende Übereinstimmung zwischen Hausumringen<br>und Gebäuden ..... | 10 |
| 3.1.3 | Defizit der Auswahlgrundlage: Existenz irrelevanter Hausumringe .....                              | 17 |
| 3.1.4 | Defizit der Auswahlgrundlage: Untererfassung .....                                                 | 25 |
| 3.1.5 | Defizit der Auswahlgrundlage: Fehlende Gebäudemerkmale .....                                       | 28 |
| 3.2   | Stichprobenauswahlverfahren .....                                                                  | 29 |
| 3.2.1 | Stichprobenumfang .....                                                                            | 29 |
| 3.2.2 | 1. Ziehungsstufe: Geschichtete Zufallsauswahl von Erhebungsbezirken .....                          | 32 |
| 3.2.3 | 2. Ziehungsstufe: Geschichtete Zufallsauswahl von Hausumringen .....                               | 41 |
| 3.2.4 | Ziehungswahrscheinlichkeiten .....                                                                 | 47 |
| 3.3   | Schätzung .....                                                                                    | 55 |
| 3.3.1 | Schätzung von Merkmalssummen .....                                                                 | 55 |
| 3.3.2 | Schätzung des Verhältnisses zweier Merkmalssummen .....                                            | 57 |
| 3.3.3 | Kombinierte Schätzung von Merkmalssummen und Gaußsches<br>Fehlerfortpflanzungsgesetz .....         | 59 |
| 4     | Nicht-stichprobenbedingte Fehler .....                                                             | 63 |
| 4.1   | Fehler infolge von Nichtbeobachtbarkeit .....                                                      | 63 |
| 4.1.1 | Undercoverage .....                                                                                | 63 |
| 4.1.2 | Nonresponse .....                                                                                  | 64 |
| 4.2   | Beobachtungsfehler .....                                                                           | 74 |
| 4.2.1 | Messfehler .....                                                                                   | 75 |
| 4.2.2 | Datenaufbereitungsfehler .....                                                                     | 75 |
| 5     | Literatur .....                                                                                    | 77 |



## 1 Zum Aufbau des Berichts

Die Stichprobenziehung sowie die sich anschließende Auswertung der erhobenen Gebäudedaten stützen sich auf ein stichprobentheoretisches Modell, das nachfolgend im Detail beschrieben und motiviert wird.

Die Darstellung beginnt mit der Vorstellung der an dieses Modell zu stellenden Anforderungen (vgl. Kapitel 2). Dabei wird sich zeigen, dass die diversen Anforderungen teilweise in einem Spannungsverhältnis zueinander stehen.

Das für die vorliegende Erhebung maßgebliche und in Kapitel 3 beschriebene Stichprobendesign stellt als Quintessenz der stichprobentheoretischen Konzeption daher einen Kompromiss dar, der versucht, weitestgehend alle in Kapitel 2 erörterten Anforderungen gemeinsam zu erfüllen. Im Zuge der Festlegung des Stichprobendesigns mussten insbesondere Entscheidungen im Hinblick auf

- die Auswahlgrundlage für die Stichprobenziehung und
- das eigentliche Stichprobenauswahlverfahren

getroffen werden, die in entsprechenden Unterkapiteln von Kapitel 3 dargelegt und begründet werden. Ziel und Zweck der Stichprobenziehung ist es, die Erhebung von Gebäudemerkmalen auf diejenigen Gebäude zu beschränken, die in die Stichprobe aufgenommen wurden, gleichzeitig aber Aussagen über die Gesamtpopulation, hier die Grundgesamtheit aller Nichtwohngebäude (mit einem besonderen Fokus auf die Teilmenge aller GEG<sup>1</sup>-relevanten Nichtwohngebäude), abzuleiten und deren Struktur und Zusammensetzung hinreichend genau zu beschreiben. Dies gelingt, indem die erhobenen Merkmale der Stichprobengebäude durch sog. Schätzverfahren in populationsbeschreibende Parameter (sog. Schätzwerte) transformiert werden. Ein solches Vorhaben ist aufgrund des Stichprobenansatzes, also des Versuchs, Aussagen über eine Grundgesamtheit auf Basis einer kleinen Teilmenge dieser Grundgesamtheit zu tätigen, notgedrungen mit Unsicherheit behaftet. Gemeinhin und so auch hier wird diese Unsicherheit durch den sog. stichprobenbedingten Fehler quantifiziert. Wie dieser im vorliegenden Fall berechnet wurde, ist ebenfalls Gegenstand dieses Unterkapitels.

Bei Befragungsprojekten wie dem vorliegenden darf eine weitere Quelle von Unsicherheit nicht außer Acht bleiben, nämlich die sog. nicht-stichprobenbedingten Fehler, die sich wiederum in Fehler infolge von Nichtbeobachtbarkeit und Beobachtungsfehler aufspalten lassen. Nicht-stichprobenbedingte Fehler und ihre Handhabung im vorliegenden Projekt stehen im Mittelpunkt des letzten Berichtskapitels (Kapitel 4).

Eine fundierte Beschreibung des stichprobentheoretischen Modells kommt ohne mathematische Formeln nicht aus. Gleichwohl wird im Folgenden auf aufwändige Herleitungen der präsentierten Formeln ebenso verzichtet wie auf Beweise und stattdessen auf die einschlägige Literatur verwiesen. Die nachfolgend verwendete Notation lehnt sich dabei grundsätzlich an die von (Särndal et al. 1992) an.

---

<sup>1</sup> Das Gebäudeenergiegesetz (GEG) trat am 1. November 2020 in Kraft und ersetzt das bisherige Energieeinsparungsgesetz (EnEG), die bisherige Energieeinsparverordnung (EnEV) und das bisherige Erneuerbare-Energien-Wärmegesetz (EEWärmeG).

Auch in der Stichprobentheorie ist das Englische die weltweit akzeptierte und genutzte Wissenschaftssprache. Dies bringt es bisweilen mit sich, dass Übersetzungen englischer Fachtermini ins Deutsche nicht eindeutig und in Einzelfällen auch missverständlich bis hin zu grob irreführend sind<sup>2</sup>. Zum Zwecke einer terminologischen Exaktheit werden in diesem Bericht daher unmittelbar nach der erstmaligen Verwendung von deutschen Fachtermini die jeweiligen englischen Pendants eingeklammert und in Kursivdruck mit angegeben.

---

<sup>2</sup> Ein Beispiel hierfür ist der bereits in diesem Einleitungskapitel verwendete deutsche Begriff „Fehler“ (*error*). Denn „Fehler“ ist nicht dahingehend zu verstehen, dass die auf Basis der Stichprobe gewonnenen Auswertungsergebnisse von den tatsächlichen Gegebenheiten in der Grundgesamt aller Nichtwohngebäude abweichen und in diesem Sinne „falsch“ sind. Stattdessen handelt es sich bei „Fehlern“ im Sinne der Stichprobentheorie um nicht gänzlich auszuräumende Unsicherheiten, die daher rühren, dass man aus unterschiedlichen Gründen nicht mit letzter Gewissheit die Übereinstimmung der Auswertungsergebnisse mit den wahren Gegebenheiten garantieren kann.

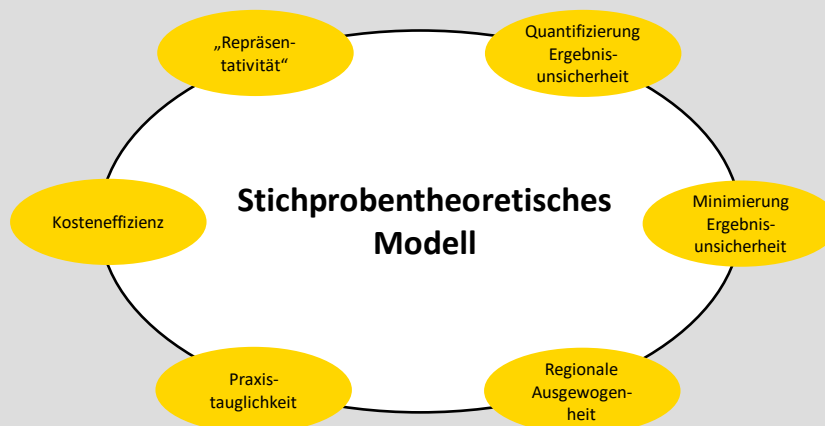
## 2 Anforderungen an das stichprobentheoretische Modell

Eine Vollerhebung (*census*) aller Nichtwohngebäude in Deutschland stand aus Kosten- und Praktikabilitätsgründen nicht zur Debatte. Aufgrund dessen stellt die statistische Modellierung der Erhebung letztendlich darauf ab, wie viele solcher Gebäude stellvertretend für den Gesamtbestand nach welchen Kriterien in die Erhebung aufgenommen werden sollen und wie die für diese Gebäude erhobenen Merkmale in Parameter (*parameter*) zu verdichten sind, die hinreichend genau die wahren Verhältnisse in der Untersuchungs- bzw. Grundgesamtheit (*(target) population*) aller Nichtwohngebäuden beschreiben<sup>3</sup>. Die statistische Modellierung konzentriert sich daher auf die zu wählende Methodik der zu ziehenden Stichprobe (*sample*) sowie auf die im Rahmen der späteren Auswertungen vorgenommene Transformation der erhobenen Merkmale der Stichprobengebäude in populationsbeschreibende Parameter.

An die Festlegung des stichprobentheoretischen Modells sind einige Anforderungen zu stellen (vgl. Abbildung 2-1). Konkret sind das im vorliegenden Fall

- „Repräsentativität“ im statistischen Sinne,
- Quantifizierbarkeit der stichprobenimmanenten Ergebnisunsicherheit,
- Minimierung der stichprobenimmanenten Ergebnisunsicherheit,
- deutschlandweit regionale Ausgewogenheit,
- Praxistauglichkeit und
- Kosteneffizienz.

**Abbildung 2-1: Anforderungen an das stichprobentheoretische Modell**



Quelle: Eigene Darstellung

<sup>3</sup> Parameter im Sinne der Statistik sind Kennziffern, die die wesentlichen Eigenschaften der Häufigkeitsverteilung eines interessierenden Merkmals beschreiben. Man unterscheidet zwischen Lageparameter (z. B. Mittelwert, Anteil) und Streuungsparameter (z. B. Varianz).

Das stichprobentheoretische Modell ist vor allem daran zu messen, inwieweit es „Repräsentativität“<sup>4</sup> herzustellen vermag. Im Sinne der Stichprobentheorie kann das nur bedeuten, dass das stichprobentheoretische Modell die Grundlage für unverzerrte (*unbiased*) bzw. erwartungstreue Ergebnisschätzungen bei möglichst geringer und quantifizierbarer Stichprobenbedingter Unsicherheit zu schaffen hat. Da ein Rückschluss von einer Stichprobe auf die interessierenden Verhältnisse der Grundgesamtheit grundsätzlich nur unter Inkaufnahme von Unsicherheit möglich ist<sup>5</sup>, d. h. der aus der gezogenen Stichprobe abgeleitete Schätzwert (*estimation value*) als Näherungs- bzw. Ersatzwert den unbekannten und deshalb zu schätzenden Parameter der Grundgesamtheit also nicht notgedrungen exakt trifft, versteht sich Erwartungstreue als ein methodisches Gütekriterium für einen solchen Rückschluss. Konkret erfordert Erwartungstreue eine methodische Vorgehensweise, die sicherstellt, dass der unbekannte und deshalb auf Stichprobenbasis zu schätzende Parameter der Grundgesamtheit (z. B. die Gesamtzahl aller Nichtwohngebäude in Deutschland oder der Anteil der Büro-, Verwaltungs- und Amtsgebäude) exakt dann getroffen wird, wenn über alle Stichproben, die nach demselben Auswahlverfahren hätten gezogen werden können<sup>6</sup>, gemittelt wird und man dabei jede dieser Stichproben mit der Wahrscheinlichkeit ihrer Realisierung gewichtet<sup>7</sup>. Das Ausmaß der mit einem solchen Rückschluss verbundenen Unsicherheit beziffert der sog. Standardfehler (*standard error*). Dieser beschreibt, wie stark ein aus der Stichprobe abgeleiteter Schätzwert von Stichprobe zu Stichprobe schwankt, und misst somit die „Interstichprobenvariabilität“. Auch wenn der Standardfehler selbst nicht auf Basis der einen gezogenen Stichprobe exakt berechnet werden kann, stehen Methoden zur Verfügung, um aus der gezogenen Stichprobe heraus einen Näherungs- bzw. Schätzwert für den Standardfehler anzugeben.

Damit insbesondere Erwartungstreue, aber auch die Quantifizierbarkeit der Stichprobenbedingten Ergebnisunsicherheit mittels des Standardfehlers möglich sind, müssen die drei folgenden Voraussetzungen erfüllt sein:

- Die Stichprobenziehung muss nach dem Zufallsprinzip erfolgen.
- Die Ziehungswahrscheinlichkeiten (*inclusion probabilities*) der in die Stichprobe aufgenommenen Nichtwohngebäude müssen bekannt, aber nicht zwingend identisch sein und bei der Schätzung (*estimation*), d. h. der

<sup>4</sup> Beim Terminus „Repräsentativität“ handelt es sich um einen nicht definierten umgangssprachlichen Begriff außerhalb der Stichprobentheorie, der deshalb hier in Anführungszeichen gesetzt und im Rahmen der nachfolgenden Ausführungen bis auf wenige Ausnahmen nicht mehr verwendet wird.

<sup>5</sup> Man spricht von einem sog. Stichprobenbedingten Fehler (*sampling error*).

<sup>6</sup> Je nach dem konkreten Auswahlverfahren kann die Zahl dieser Stichproben unvorstellbar groß sein. Selbstverständlich wird in der Praxis nur eine einzige Stichprobe gezogen, anhand derer Rückschlüsse bezüglich der Grundgesamtheit gezogen werden müssen.

<sup>7</sup> Zu beachten ist, dass Erwartungstreue keine Eigenschaft der Stichprobenziehung selbst, sondern der sich anschließenden Schätzung (*estimation*) ist. Erwartungstreue ist daher eine Eigenschaft der Vorgehensweise, durch die Stichprobenergebnisse auf die Grundgesamtheit übertragen werden. Beim auf Basis der einen gezogenen Stichprobe schlussendlich ermittelten Ersatzwert für den interessierenden Populationsparameter handelt es sich um einen Schätzwert (*estimation value*), der auch über oder unter dem zu schätzenden Parameter der Grundgesamtheit liegen kann, ohne dass man die Abweichung zum zu schätzenden Parameter der Grundgesamtheit beziffern kann (denn anderweitig würde man den interessierenden Populationsparameter ja kennen und eine Schätzung auf Stichprobenbasis wäre redundant).



Übertragung von Stichprobenergebnissen auf die Grundgesamtheit, adäquat berücksichtigt werden.

- Alle Nichtwohngebäude müssen eine Chance haben, in die Stichprobe aufgenommen zu werden. Sie müssen also positive Ziehungswahrscheinlichkeiten aufweisen.

Darüber hinaus gewährleistet die nachfolgende vierte Bedingung, dass (approximativ) erwartungstreue Schätzungen der Standardfehler vorgenommen werden können<sup>8</sup>:

- Die Ziehungswahrscheinlichkeiten dafür, dass zwei beliebige Nichtwohngebäude gemeinsam in die Stichprobe aufgenommen werden, müssen positiv und bekannt sein.

Auch wenn die Ergebnisunsicherheit stichprobenimmanent ist, so sollte sie doch möglichst gering ausfallen. Ein zentraler Ansatzpunkt dafür ist der Stichprobenumfang (*sample size*), der unter gleichen Bedingungen so hoch wie möglich angesetzt werden sollte. Daneben hält die Stichprobentheorie noch eine Reihe weiterer methodischer Instrumente vor, die (unter anderem) auf eine Minimierung der Ergebnisunsicherheit abzielen.

Das stichprobentheoretische Modell hat im vorliegenden Fall auch eine regionale Ausgewogenheit der Stichprobe zu gewährleisten, d. h. es muss dafür Sorge tragen, dass ausreichend viele Nichtwohngebäude aus allen Teilen Deutschlands in der Stichprobe vertreten sind, denn nur dann sind belastbare Auswertungen unterhalb der Bundesebene möglich.

Dass die Stichprobenmethodik praxistauglich sein soll und gleichzeitig zu vertretbaren Kosten umsetzbar sein muss, sind zwei weitere Anforderungen.

Zu beachten ist, dass die beschriebenen sechs Anforderungen teilweise in Konflikt zueinander stehen, wodurch nicht alle Anforderungen gleichzeitig in vollem Umfang zu erfüllen sind. So erfordert die gebotene Begrenzung der Erhebungskosten einen möglichst kleinen Stichprobenumfang, was jedoch dem Ziel zuwiderläuft, die stichprobenbedingte Ergebnisunsicherheit möglichst gering zu halten. Aufgrund der allgegenwärtigen Existenz solcher Zielkonflikte ist die Festlegung eines Stichprobendesigns stets und somit auch im vorliegenden Fall das Ergebnis eines Abwägungsprozesses zwischen den gestellten Anforderungen.

---

<sup>8</sup> Während der quadrierte Standardfehler im Fall von Merkmalssummenschätzungen bei Verwendung geeigneter und auch hier eingesetzter Formeln erwartungstreu geschätzt werden kann, kann bei der Schätzung des Standardfehlers selbst nur eine approximative, also näherungsweise Erwartungstreue erreicht werden (vgl. auch Unterkapitel 3.3.1).

### 3 Stichprobendesign und Schätzung

Beim Stichprobendesign (*sampling design*) handelt es sich um ein mathematisches Konstrukt, das im vorliegenden Fall vereinfacht gesprochen angibt, wie viele Nichtwohngebäude nach welchen Kriterien in die Stichprobe aufgenommen werden<sup>9</sup>. Das Stichprobendesign bestimmt dadurch schlussendlich, mit welcher Wahrscheinlichkeit ein beliebiges Nichtwohngebäude in die Stichprobe gelangt. Das Stichprobendesign ergibt sich durch die Implementierung eines geeigneten Stichprobenauswahl- bzw. -ziehungsverfahrens (*sample selection scheme, sampling scheme*) (Särndal et al. 1992 S. 27).

Ein Stichprobenauswahlverfahren stützt sich in aller Regel und so auch hier auf eine geeignete Auswahlgrundlage (*sampling frame*). Dabei handelt es sich um eine – möglichst in Gestalt einer Datenbank aufbereitete – Zusammenstellung von Elementen, konkret sog. Erhebungseinheiten (*sampling units*), die einen Zugang zu den eigentlich interessierenden Untersuchungseinheiten (*population units*), hier den Nichtwohngebäuden in Deutschland, herstellen können. Im Rahmen der stichprobentheoretischen Modellierung ist eine begründete Festlegung einer Auswahlgrundlage vorzunehmen (vgl. Unterkapitel 3.1). Nach der Ziehung der Stichprobe aus dieser Auswahlgrundlage (vgl. Unterkapitel 3.2) und dem anschließenden Erheben der interessierenden Gebäudemerkmale steht die Gewichtung bzw. Hochrechnung dieser Stichprobenfälle durch den Einsatz geeigneter Schätzverfahren an, um von der Stichprobe auf die übergeordnete Grundgesamtheit aller in Deutschland stehender Nichtwohngebäude schließen zu können (vgl. Unterkapitel 3.3). Die mit der Schätzung einhergehende stichprobenbedingte Unsicherheit wird dabei durch den Standardfehler abgebildet. Gewichtung bzw. Hochrechnung sowie Fehlerrechnung stehen in enger Beziehung zum durch das Stichprobenauswahlverfahren festgelegten Stichprobendesign.

#### 3.1 Auswahlgrundlage

An eine Auswahlgrundlage sind zwei Minimalanforderungen zu stellen. Erstens müssen die in der Auswahlgrundlage enthaltenen Erhebungseinheiten in Verbindung zu (möglichst) allen Untersuchungseinheiten, im vorliegenden Fall also zu allen Nichtwohngebäuden, stehen; denn nur dann lässt sich die in Kapitel 2 erläuterte Voraussetzung für Erwartungstreue und Quantifizierbarkeit der stichprobenbedingten Ergebnisunsicherheit mittels des Standardfehlers erfüllen, nämlich die Einräumung positiver Ziehungswahrscheinlichkeiten für alle Nichtwohngebäude. Zweitens muss sichergestellt sein, dass diejenigen Untersuchungseinheiten bzw. Nichtwohngebäude, die zu gezogenen Erhebungseinheiten „gehören“, „kontaktierbar“ sind, dass also die gewünschten Informationen über sie erhoben werden können. Voraussetzung für die Kontaktierbarkeit ist dabei die eindeutige Identifizierbarkeit der Nichtwohngebäude.

Neben diesen Minimalanforderungen erfüllt eine ideale Auswahlgrundlage in Abhängigkeit von Untersuchungsgegenstand und -ziel noch weitere Anforderungen. Bezogen auf die vorliegende Untersuchung wäre die ideale Auswahlgrundlage

<sup>9</sup> Mathematisch ausgedrückt, legt das Stichprobendesign die auf dem Stichprobenraum definierte Wahrscheinlichkeitsverteilung und damit die Wahrscheinlichkeit für das Ziehen einer bestimmten Stichprobe fest (Särndal et al. 1992 S. 27).

eine aktuelle Liste oder Datenbank aller (und nur dieser) Nichtwohngebäude in Deutschland mitsamt Gebäudeanschrift und möglichst angereichert mit weiteren Informationen zu grundlegenden Gebäudemerkmalen (insbesondere der GEG-Relevanz) sowie zu geeigneten Gebäudeansprechpartnern. Abbildung 3-1 illustriert beispielhaft eine solche ideale Auswahlgrundlage.

**Abbildung 3-1: Beispielhafter Aufbau einer idealen Auswahlgrundlage**

| Gebäudeanschrift                    | Gebäudemerkmale                                      | Ansprechpartner                       |
|-------------------------------------|------------------------------------------------------|---------------------------------------|
| Musterstraße 5<br>12345 Musterstadt | Schule, Baujahr 1980,<br>Nettoraumfläche 1000 qm     | Max Mustermann<br>Tel.: 01234/56789   |
| Musterweg 17<br>54321 Musterdorf    | Supermarkt, Baujahr 2010,<br>Nettoraumfläche 2000 qm | Marta Musterfrau<br>Tel.: 09876/54321 |
| Musterpfad 6<br>99999 Musterstadt   | Bürogebäude, Baujahr 1975,<br>Nettoraumfläche 800 qm | Musterbau AG<br>Tel.: 09999/99999     |

Quelle: Eigene Darstellung

Bei einer solchen Auswahlgrundlage könnte mit dem Wissen um geeignete Gebäudeansprechpartner (rechte Tabellenspalte) unmittelbar nach der Stichprobenziehung mit der Führung der (Telefon-) Interviews gestartet werden. Darüber hinaus würden die Untersuchungseinheiten der vorliegenden Untersuchung, also die Nichtwohngebäude, deren Adressen in der linken Tabellenspalte aufgeführt sind, gleichzeitig auch als Erhebungseinheiten fungieren, was die Ziehung der Stichprobe wie auch die spätere Übertragung der Stichprobenbefunde auf die Grundgesamtheit deutlich erleichtern würde. Dank der Kenntnis grundlegender Gebäudemerkmale (mittlere Tabellenspalte) könnte man für die Grundgesamtheit aller Nichtwohngebäude überdies entsprechende Strukturinformationen direkt aus der Datenbank und ganz ohne jede stichprobenbedingte Unsicherheit ableiten, z. B. die Anzahl aller GEG-relevanten Nichtwohngebäude oder die Anzahl aller vor der Jahrtausendwende errichteten Nichtwohngebäude. Darüber hinaus könnten diese Informationen genutzt werden, um die Schätzgenauigkeit derjenigen Parameter zu erhöhen, zu denen die Datenbank keine Informationen bereithält und die daher nur auf Stichprobenbasis bestimmt werden können<sup>10</sup>. Und Gebäudeausfälle infolge fehlgeschlagener Interviews könnten dank der in der Datenbank enthaltenen Gebäudemerkmale nicht nur auf Selektivität geprüft werden, sondern bei nachgewiesener Selektivität auch durch geeignete stichprobenmethodische Maßnahmen dahingehend korrigiert werden, dass etwaige Verzerrungen bei der Schätzung von Populationsparametern zurückgedrängt werden.

Eine Auswahlgrundlage der geschilderten Art gibt es für Nichtwohngebäude jedoch nicht und ihre Erstellung wäre mit immensen Kosten verbunden. Ersatzweise

<sup>10</sup> Die vorhandenen Informationen können sowohl bei der Ziehung der Stichprobe als auch im Rahmen der Schätzung genutzt werden.

wurde im vorliegenden Projekt auf eine vom Verbundpartner IÖR aufbereitete<sup>11</sup> Datenbank sog. amtlicher Hausumringe Deutschland (HU-DE) zurückgegriffen, d. h. auf eine Datenbank georeferenzierter Umringspolygone, abgeleitet aus der amtlichen Liegenschaftskarte zur Beschreibung von Gebäudegrundrissen. Die angesprochene Aufbereitung der amtlichen Hausumringe durch das IÖR zielte zum einen darauf ab, doppelt vorhandene Hausumringe, Kleinstpolygone, bei denen es sich nicht um eigenständige Nichtwohngebäude handeln kann, Überlappungen und atypische Grundrisse wahlweise zu beseitigen oder mit anderen Hausumringen so zu verschmelzen, dass eine größere Kongruenz zwischen Hausumring und tatsächlichem Gebäudegrundriss erreicht wird, denn entgegen der Erwartung umschreiben amtliche Hausumringe nicht immer die tatsächlichen Gebäudegrundrisse. Durch die Aufbereitung reduzierte sich der Anfangsbestand von 51.808.048 amtlichen Hausumringen auf 48.791.300 aufbereitete Hausumringe<sup>12</sup>. Zum anderen reicherte das IÖR die Hausumringe mit numerischen und semantischen Merkmalen unter Rückgriff auf diverse Datenquellen mit dem Ziel an, Rückschlüsse auf die Untersuchungsrelevanz der durch die Hausumringe beschriebenen Gebäude mit besonderem Fokus auf die GEG-Relevanz zuzulassen. Darüber hinaus fügte das IÖR den Hausumringen – soweit verfügbar – Adressinformationen zu.

Die Hausumring-Datenbank erfüllt zwar die eingangs beschriebenen Minimalanforderungen an eine Auswahlgrundlage. Doch trotz der umfangreichen und anspruchsvollen Aufbereitungsmaßnahmen ist die Hausumring-Datenbank keine ideale Auswahlgrundlage im beschriebenen Sinn. Gründe hierfür sind:

- Informationen zu Gebäudeansprechpartnern fehlen nach wie vor. Auch konnte nicht allen Hausumringen im Zuge der Aufbereitung eine Adresse zugeordnet werden. Wenn nun schon Informationen zu Gebäudeansprechpartnern fehlen, ist die genaue Kenntnis der Gebäudeadresse aus erhebungspraktischen Gründen aber mitentscheidend, um auf diesem Wege potentielle Gebäudeansprechpartner recherchieren zu können.
- Eine Deckungsgleichheit zwischen den durch die Hausumringe beschriebenen Flächen und den tatsächlichen Gebäudegrundrissen ist auch nach der Aufbereitung der Datenbank nach wie vor nicht gesichert. Es besteht somit nicht notwendigerweise eine wechselseitig eindeutige Beziehung zwischen Hausumring und Nichtwohngebäude, wobei ex ante nicht bekannt ist, welches Beziehungsgeflecht im Einzelfall konkret vorliegt. Die Kenntnis des Beziehungsgeflechtes ist jedoch erforderlich, um den letztlich interessierenden Nichtwohngebäuden die richtigen Ziehungswahrscheinlichkeiten zuordnen zu können.
- Es muss davon ausgegangen werden, dass nur ein kleiner Teil der Hausumringe zu Nichtwohngebäuden im Allgemeinen und zu GEG-relevanten Nichtwohngebäuden im Besonderen gehört. Allerdings ist eine Relevanzfeststellung auf Basis der vorhandenen hausumringspezifischen Nutzungsattribute (sog. LoD1-Gebäudefunktionen) mit Unsicherheit behaftet. Insbesondere ist nicht auszuschließen, dass durch zwischenzeitliche

<sup>11</sup> Im Folgenden wird die vom IÖR geleistete Aufbereitung nur sehr kurz umschrieben. Detaillierte Informationen finden sich in (Hartmann et al. 2020).

<sup>12</sup> Wenn nachfolgend von Hausumringen die Rede ist, sind stets die durch das IÖR aufbereiteten Hausumringe gemeint, da diese (und nicht die „ursprünglichen“ amtlichen Hausumringe) als Erhebungseinheiten fungieren und folglich die Auswahlgrundlage bilden.

Umnutzung ein Gebäude entweder seine Untersuchungsrelevanz verloren hat oder aber erst untersuchungsrelevant geworden ist. Abgesehen davon liegen von fast 20 Mio. der knapp 49 Mio. Hausumringe keine bzw. keine ausreichend aussagekräftigen Nutzungsattribute vor, anhand derer direkt auf die Untersuchungsrelevanz geschlossen werden kann.

- Generell stellt sich die Frage, ob die Datenbank auch wirklich zu allen Nichtwohngebäuden in Deutschland Hausumringe bereitstellt. Ganz konkret stellt sich in diesem Zusammenhang die Frage nach der Aktualität der herangezogenen Datenbank amtlicher Hausumringe Deutschland, denn der Datenbankauszug datiert vom April 2015. Gebäude jüngeren Baualters sind daher per se nicht über die Hausumring-Datenbank erreichbar.<sup>13</sup>
- Gebäudemerkmale der gewünschten Art liegen bis auf die Nutzungsattribute nicht vor, und wie ausgeführt handelt es sich bei Letzteren nicht um gebäude-, sondern um hausumringbezogene Merkmale, die ggf. nicht mehr aktuell sind und überdies nur für einen Teil der Hausumringe vorliegen.

Die aufgelisteten Defizite der Auswahlgrundlage erfordern sich wechselseitig beeinflussende erhebungspraktische wie stichprobenmethodische Maßnahmen, die nachfolgend erörtert werden.

### 3.1.1 Defizit der Auswahlgrundlage: Fehlende Informationen zu Gebäudeansprechpartnern

Dank der Georeferenzierung der Hausumringe in Gestalt geografischer Koordinaten ist zwar deren Verortung im Raum gewährleistet und dadurch jedes Nichtwohngebäude, das mit Hausumringen Schnittflächen aufweist, identifizierbar. Die Erhebung der gewünschten gebäudespezifischen Informationen bedingt jedoch die Kontaktierung eines geeigneten, also kompetenten und auskunftswilligen Ansprechpartners, idealerweise des Gebäudeeigentümers, ersatzweise auch eines Gebäudenutzers (sofern Gebäudeeigentümer und -nutzer auseinanderfallen). Da die Nutzung von im Liegenschaftskataster hinterlegten Eigentümerinformationen jedoch aus datenschutzrechtlichen Gründen nicht möglich war, mussten geeignete Gebäudeansprechpartner auf anderem Wege recherchiert werden. Eine (Internet-) Recherche von Gebäudenutzern erfordert dabei Hintergrundinformationen, etwa die Kenntnis der postalischen Gebäudeadresse. Wie beschrieben, liegen aber selbst nach der Aufbereitung der Hausumringdaten nicht für alle Hausumringe Adressinformationen vor. Darüber hinaus ist aufgrund der nicht notwendigerweise wechselseitig eindeutigen Beziehung zwischen Hausumring und Nichtwohngebäude nicht gewährleistet, dass eine vorliegende Hausumringadresse auch der Gebäudeadresse bzw. – im Fall mehrerer dem Hausumring zugeordneter Gebäude – den Gebäudeadressen entspricht. Aber auch abgesehen davon reicht die Kenntnis einer Gebäudeadresse für sich genommen in vielen Fällen nicht aus, um darauf aufbauend Kontaktinformationen zu Gebäudenutzern zu recherchieren. Valide Informationen zu Gebäudenutzern sind in letzter Konsequenz nur vor Ort recherchierbar, indem – unter Wahrung datenschutzrechtlicher Vorgaben

<sup>13</sup> In Ausnahmefällen können auch Gebäude jüngeren Baualters über die Hausumring-Datenbank erreichbar sein, nämlich dann, wenn sie auf Flächen inzwischen abgerissener Gebäude stehen, deren Hausumring(e) im Datenbankauszug enthalten sind (vgl. Unterkapitel 3.1.4).

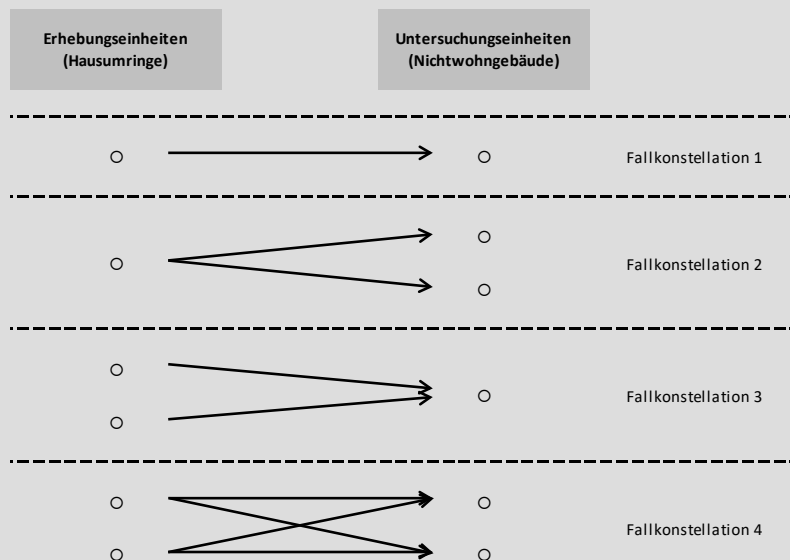
vom öffentlichen Raum aus – Firmenlogos oder Klingelschilder aufgezeichnet werden. Kurzum, zur Recherche von Gebäudeansprechpartnern ist eine Vor-Ort-Begehung („Screening“) unausweichlich. Die Einrichtung eines solchen Screenings ist dabei eine aufwändige erhebungspraktische Maßnahme mit großen Auswirkungen auf das Stichprobenauswahlverfahren und daher auf die gesamte stichprobentheoretische Modellierung der Erhebung. Wie bei der nachfolgenden Besprechung der anderen Defizite der vorliegenden Auswahlgrundlage verdeutlicht wird, gibt es für ein solches Screening weitere Argumente; die zentrale Notwendigkeit ergibt sich allerdings aus dem Umstand, dass es ohne eine solche Vor-Ort-Begehung nicht lückenlos möglich ist, an Informationen zu gelangen, die die Identifizierung geeigneter Gebäudeansprechpartner erlauben.

### **3.1.2 Defizit der Auswahlgrundlage: Fehlende Übereinstimmung zwischen Hausumringen und Gebäuden**

Wie bereits zu Beginn von Kapitel 3 ausgeführt, ist zwischen Untersuchungs- und Erhebungseinheiten zu unterscheiden. Untersuchungseinheiten sind im vorliegenden Fall Nichtwohngebäude entsprechend der hier angelegten Definition. Während die Untersuchungseinheiten durch das inhaltliche Erkenntnisinteresse der Erhebung festgelegt sind und daher denjenigen Einheiten entsprechen, über die im Einklang mit dem Untersuchungsziel am Ende „repräsentative“ datenorientierte Aussagen getroffen werden sollen, erstreckt sich die eigentliche Stichprobenziehung auf Erhebungseinheiten (Stenger 1986 S. 9), d. h. im vorliegenden Fall auf Hausumringe.

Wie ebenfalls bereits an früherer Stelle konstatiert, besteht ein – behebbares – Defizit der vorliegenden Auswahlgrundlage darin, dass nicht von vornherein von einer wechselseitig eindeutigen, also einer eineindeutigen (1:1-) Beziehung zwischen der Erhebungseinheit „Hausumring“ und der Untersuchungseinheit „Nichtwohngebäude“ ausgegangen werden kann. Es ist also nicht davon auszugehen, dass auf jeder durch einen Hausumring umschriebenen Fläche stets nur ein einziges Nichtwohngebäude (ganz oder teilweise) steht und dieses Nichtwohngebäude gleichzeitig keinerlei Berührungspunkte mit anderen Hausumringen hat. Vielmehr sind insgesamt vier verschiedene Fallkonstellationen denkbar, die in allgemeingültiger Form in Abbildung 3-2 veranschaulicht werden.

**Abbildung 3-2: Mögliche Beziehungsgeflechte zwischen Erhebungseinheiten (Hausumringe) und Untersuchungseinheiten (Nichtwohngebäude)**

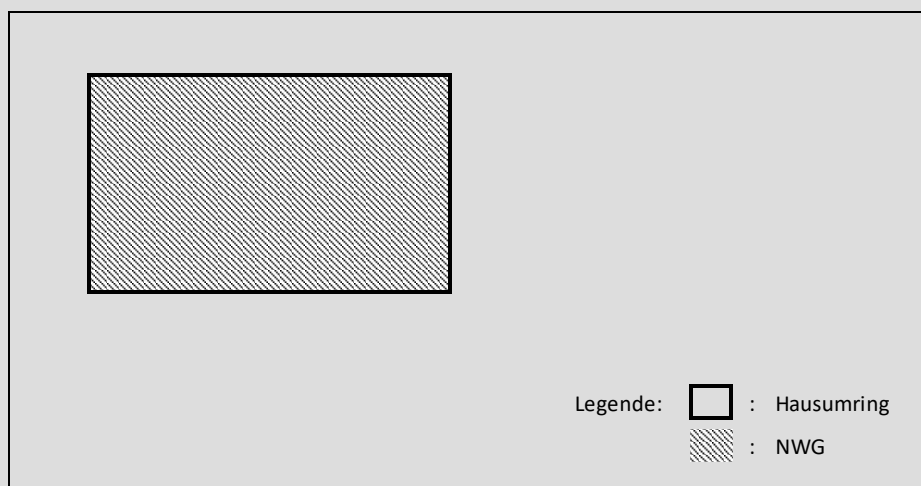


Quelle: Eigene Darstellung

Wie an späterer Stelle (in Unterkapitel 3.2.4) noch erläutert wird, ist die Kenntnis der jeweiligen Fallkonstellation notwendig, um den Nichtwohngebäuden die richtigen Ziehungswahrscheinlichkeiten zuzuordnen.

**Fallkonstellation 1** ist der einfachste Fall einer eindeutigen Beziehung zwischen Erhebungs- und Untersuchungseinheit. Die Erhebungseinheit zeigt alleine und ausschließlich auf eine einzige Untersuchungseinheit. Bezogen auf die vorliegende Erhebung kann, muss dabei aber nicht die in Abbildung 3-3 dargestellte Situation einer flächenmäßigen Deckungsgleichheit zwischen Hausumringfläche und Gebäudegrundriss gegeben sein.

**Abbildung 3-3: Mögliche Beziehungsgeflechte zwischen Erhebungseinheiten (Hausumringe) und Untersuchungseinheiten (Nichtwohngebäude): Fallkonstellation 1, Beispiel 1**



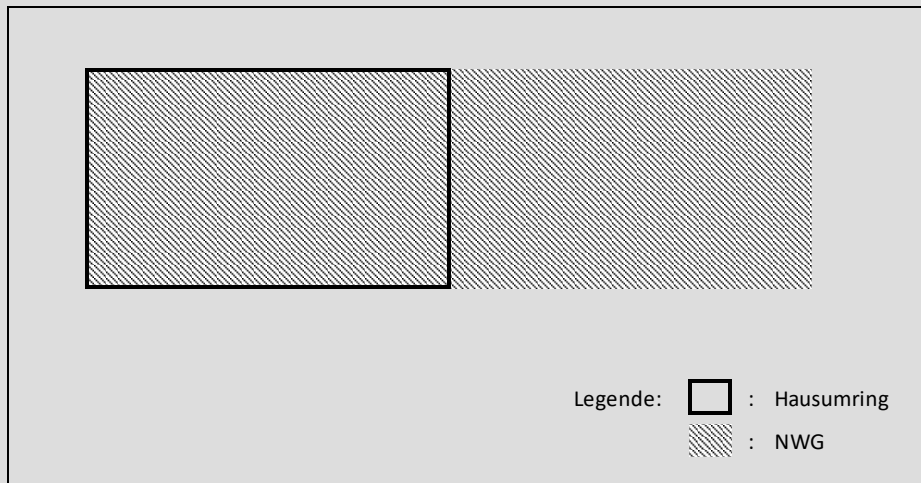
Quelle: Eigene Darstellung

Denn auch wenn die Fläche des Hausumrings nicht exakt mit dem Grundriss des Nichtwohngebäudes zusammenfällt, kann immer noch Fallkonstellation 1 vorliegen, nämlich dann, wenn sich die Fläche des Hausumrings und der Gebäudegrundriss schneiden<sup>14</sup>, aber weder der Hausumring den Grundriss eines anderen Nichtwohngebäudes (zumindest partiell) überdeckt, noch das betreffende Gebäude Schnittflächen mit anderen Hausumringen aufweist. In Abbildung 3-4 wird ein solches Beispiel grafisch dargestellt. Die Hausumringfläche ist im Beispiel deutlich kleiner als der Gebäudegrundriss, was möglicherweise darauf zurückzuführen ist, dass das Gebäude im Laufe der Zeit nach rechts erweitert wurde, ohne dass diese Erweiterung – aus welchen Gründen auch immer – Niederschlag im Liegenschaftskataster gefunden hat, d. h. ohne dass die Geometrie des Hausumrings entsprechend angepasst oder ein weiterer Hausumring gebildet wurde. Gleichwohl liegt weiterhin Fallkonstellation 1 vor, da der Hausumring als einziger und ausschließlich auf das betreffende Nichtwohngebäude zeigt, d. h. kein weiteres Nichtwohngebäude überdeckt, und gleichzeitig das von diesem Hausumring (partiell) überdeckte Gebäude nicht durch weitere Hausumringe überdeckt wird.

<sup>14</sup> Mengentheoretisch ausgedrückt darf die Schnittmenge nicht leer sein.



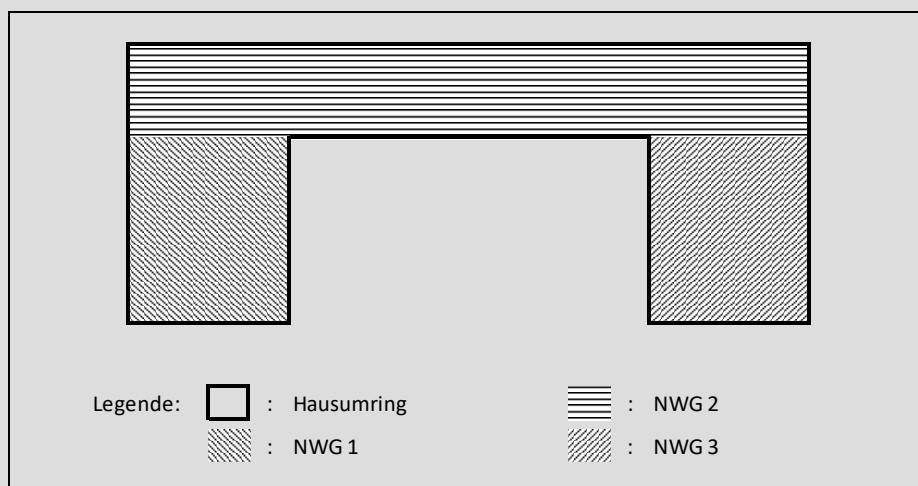
**Abbildung 3-4:** Mögliche Beziehungsgeflechte zwischen Erhebungseinheiten (Hausumringe) und Untersuchungseinheiten (Nichtwohngebäude): Fallkonstellation 1, Beispiel 2



Quelle: Eigene Darstellung

Bei **Fallkonstellation 2** zeigt eine Erhebungseinheit auf mehr als eine Untersuchungseinheit. Umgekehrt zeigt auf jede dieser Untersuchungseinheiten jedoch stets nur eine und zwar dieselbe Erhebungseinheit. Fallkonstellation 2 ist daher immer dann gegeben, wenn – wie im Beispiel der Abbildung 3-5 – ein Hausumring mehrere Gebäudegrundrisse umschließt. Die in der Abbildung gegebene vollständige Deckungsgleichheit zwischen der Fläche des hufeisenförmigen Hausumrings auf der einen und den drei Gebäudegrundrissen auf der anderen Seite stellt dabei keine notwendige Bedingung für Fallkonstellation 2 dar. Denn auch wenn beispielsweise das Nichtwohngebäude 3 über die Fläche des Hausumrings hinausragt und somit nur teilweise durch den Hausumring überdeckt werden würde oder aber wenn umgekehrt der Hausumring über Nichtwohngebäude 3 hinaus in den unbebauten Raum ragen würde, läge immer noch Fallkonstellation 2 vor.

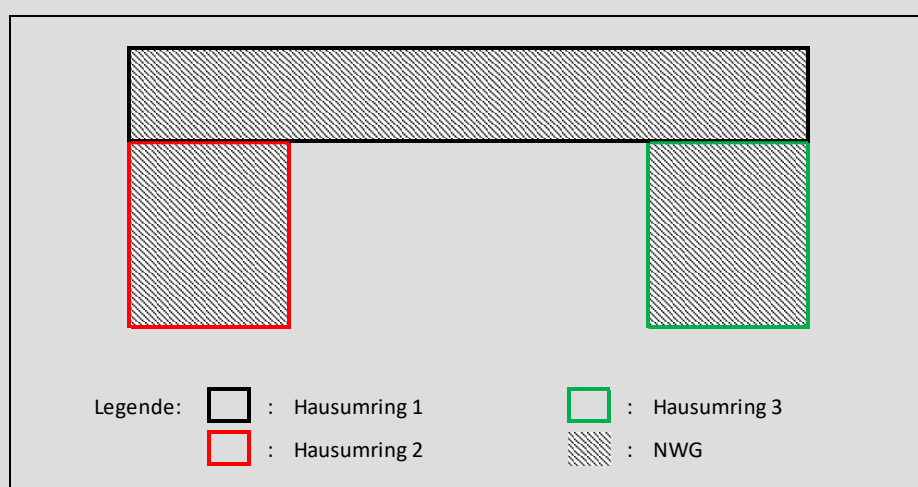
**Abbildung 3-5: Mögliche Beziehungsgeflechte zwischen Erhebungseinheiten (Hausumringe) und Untersuchungseinheiten (Nichtwohngebäude): Fallkonstellation 2, Beispiel**



Quelle: Eigene Darstellung

**Fallkonstellation 3** ist in gewisser Weise das Gegenstück zu Fallkonstellation 2, denn jetzt zeigen mehrere Erhebungseinheiten auf dieselbe Untersuchungseinheit. Abbildung 3-6 illustriert einen Beispielfall, bei dem sich die Grundfläche des Nichtwohngebäudes additiv aus drei Hausumringflächen zusammensetzt.

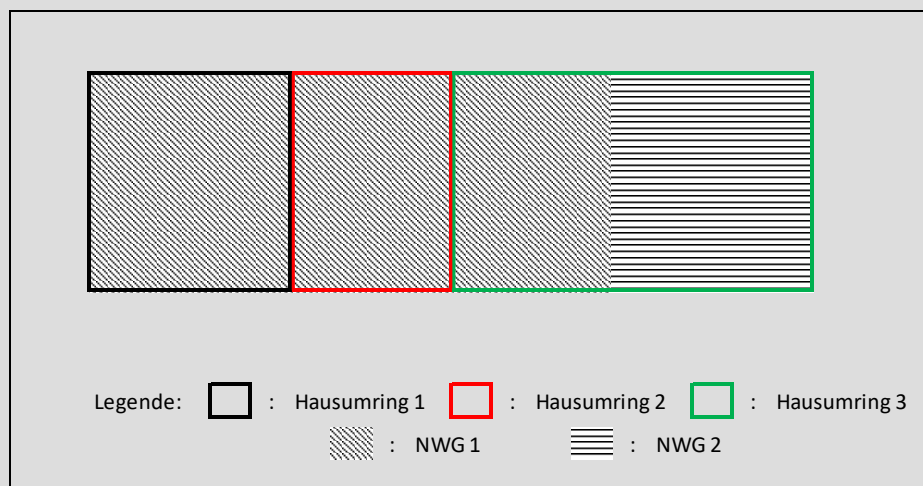
**Abbildung 3-6: Mögliche Beziehungsgeflechte zwischen Erhebungseinheiten (Hausumringe) und Untersuchungseinheiten (Nichtwohngebäude): Fallkonstellation 3, Beispiel 1**



Quelle: Eigene Darstellung

Im Gegensatz zu den bisher betrachteten Fallkonstellationen 1 und 2, bei denen stets nur eine einzige Erhebungseinheit vorlag, bedingt die Feststellung von Fallkonstellation 3 allerdings die vorherige Festlegung der zu betrachtenden Erhebungseinheit. In Abbildung 3-6 führt diese Festlegung aufgrund der völligen Deckungsgleichheit zwischen den zusammengesetzten drei Hausumringflächen und dem Gebäudegrundriss zwar stets zum selben Ergebnis („Fallkonstellation 3“), Abbildung 3-7 illustriert aber einen Beispielfall, in dem die Feststellung von Fallkonstellation 3 nur in Bezug auf die Hausumringe 1 und 2, nicht aber in Bezug auf Hausumring 3 zutrifft: Die Fläche von Hausumring 1 hat nur mit einem einzigen Nichtwohngebäude, nämlich Nichtwohngebäude 1, eine Schnittfläche, wobei dieses Nichtwohngebäude auch durch zwei weitere Hausumringe, nämlich die Hausumringe 2 und 3, überdeckt wird. In Bezug auf Hausumring 1 liegt folglich Fallkonstellation 3 vor. Gleiches gilt in Bezug auf Hausumring 2. In Bezug auf Hausumring 3 lässt sich die Feststellung von Fallkonstellation 3 dagegen nicht aufrechterhalten, da Hausumring 3 eine gemeinsame Fläche mit einem weiteren Nichtwohngebäude, hier mit Nichtwohngebäude 2, hat.

**Abbildung 3-7:** Mögliche Beziehungsgeflechte zwischen Erhebungseinheiten (Hausumringe) und Untersuchungseinheiten (Nichtwohngebäude): Fallkonstellation 3 (für Hausumring 1 und 2), Beispiel 2; Fallkonstellation 4 (für Hausumring 3), Beispiel 1



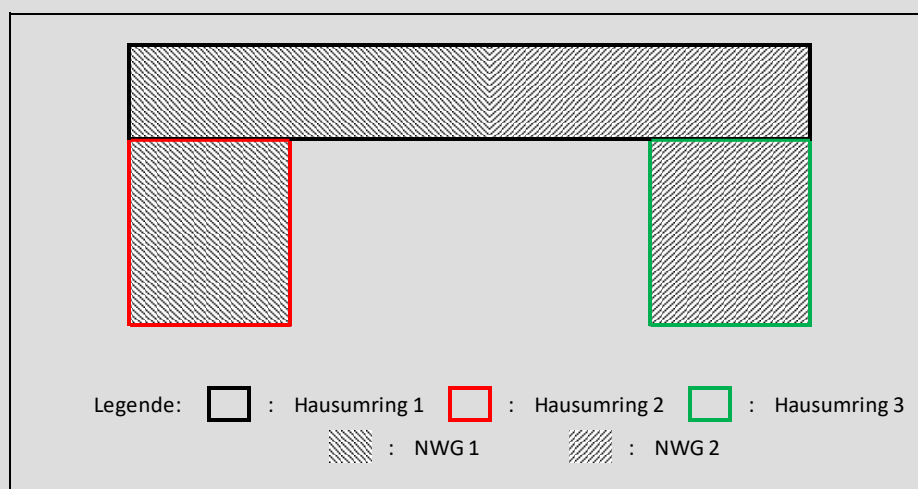
Quelle: Eigene Darstellung

Für Hausumring 3 stellt sich daher eine Situation dar, die durch keine der drei bisher betrachteten Fallkonstellationen 1 bis 3 für sich genommen vollumfänglich beschrieben wird. Vielmehr liegt eine Art Mischung aus den beiden Fallkonstellationen 2 und 3 vor, da zum einen Hausumring 3 die Grundrisse zweier Nichtwohngebäude (ganz oder teilweise) überdeckt (Fallkonstellation 2) und da sich zum anderen unter diesen beiden Nichtwohngebäuden mit Nichtwohngebäude 1 eines befindet, dessen Grundfläche auch die Flächen anderer Hausumringe (hier:

Hausumringe 1 und 2) schneidet (Fallkonstellation 3). Solche gemischten Fallkonstellationen definieren eine neue Fallkonstellation, nämlich **Fallkonstellation 4**. Fallkonstellation 4 ist – in methodischer Terminologie ausgedrückt – somit dadurch gekennzeichnet, dass die betrachtete Erhebungseinheit auf mehr als eine Untersuchungseinheit zeigt und dass sich unter diesen Untersuchungseinheiten wenigstens eine befindet, auf die zumindest eine weitere Erhebungseinheit zeigt.

Abbildung 3-8 veranschaulicht ein weiteres Beispiel für Fallkonstellation 4. Die Fläche von Hausumring 1 schneidet dort zwei Gebäudegrundflächen, die jedoch auch von anderen Hausumringen partiell überdeckt werden. Konkret ist das bei Nichtwohngebäude 1 Hausumring 2 und bei Nichtwohngebäude 2 Hausumring 3. In Bezug auf Hausumring 1 liegt somit Fallkonstellation 4 vor. Anders verhält es sich dagegen in Bezug auf die beiden anderen Hausumringe, für die jeweils Fallkonstellation 3 zutrifft.

**Abbildung 3-8:** Mögliche Beziehungsgeflechte zwischen Erhebungseinheiten (Hausumringe) und Untersuchungseinheiten (Nichtwohngebäude): Fallkonstellation 4 (für Hausumring 1), Beispiel 2



Quelle: Eigene Darstellung

Tabelle 3-1 fasst die vier möglichen Fallkonstellationen zwischen dem betrachteten Hausumring und den durch ihn vollständig oder partiell abgedeckten Nichtwohngebäuden zusammen.

**Tabelle 3-1: Tabellarische Darstellung der vier möglichen Hausumring-NWG-Fallkonstellationen**

|                                                   |                        | NWG hat/haben Schnittfläche(n) mit ... |                       |
|---------------------------------------------------|------------------------|----------------------------------------|-----------------------|
|                                                   |                        | ... einem einzigen HU                  | ... mehr als einem HU |
| Betrachteter Hausumring hat Schnittfläche mit ... | ... einem einzigen NWG | Fallkonstellation 1                    | Fallkonstellation 3   |
|                                                   | ... mehr als einem NWG | Fallkonstellation 2                    | Fallkonstellation 4   |

Quelle: Eigene Darstellung

Die Feststellung der jeweils zutreffenden Fallkonstellation und damit letztlich die Klärung der konkreten Bebauungssituation wäre prinzipiell auch im Austausch mit den recherchierten Gebäudeansprechpartnern möglich gewesen. Wenn nun aber ohnehin schon aus anderem Grund (Einholung von Informationen zur Recherche geeigneter Gebäudeansprechpartner, vgl. Unterkapitel 3.1.1) eine Vor-Ort-Begehung erforderlich ist, drängt es sich auf, die Aufgabe der Klärung des Beziehungsgeflechts zwischen dem jeweiligen Stichproben-Hausumring und der auf seiner Fläche vorzufindenden Bebauung mit Nichtwohngebäuden dem Screening-Personal zu übertragen. Dies bringt nämlich zwei Vorteile mit sich. Zum einen kann durch die entsprechende Rekrutierung und Schulung des Screening-Personals dafür Sorge getragen werden, dass die im Rahmen dieses Projekts festgelegte Nichtwohngebäude-Definition adäquat umgesetzt und die jeweils zutreffende Fallkonstellation richtig erkannt wird. Zum anderen besteht nicht das Risiko einer ungeklärten Bebauungssituation, wenn ein geeigneter Gebäudeansprechpartner nicht gefunden werden kann oder wenn dieser Angaben verweigert. Indem die Aufgabe der Klärung des Beziehungsgeflechts zwischen Stichproben-Hausumring und Nichtwohngebäude(n) dem Screening-Personal übertragen wird, kann vielmehr sichergestellt werden, dass – mit Ausnahme nicht-einsehbarer Bebauungssituationen – für jeden Stichproben-Hausumring am Ende der Screeningphase feststeht, mit welchen Nichtwohngebäuden er gemeinsame Flächen aufweist. In Verbindung mit der Gewährleistung der sachgerechten Umsetzung der Nichtwohngebäudedefinition kann dadurch der unbekannte Umfang der Grundgesamtheit, also die Anzahl der in Deutschland stehenden Nichtwohngebäude, verzerrungsfrei geschätzt werden (vgl. Unterkapitel 3.3.1).

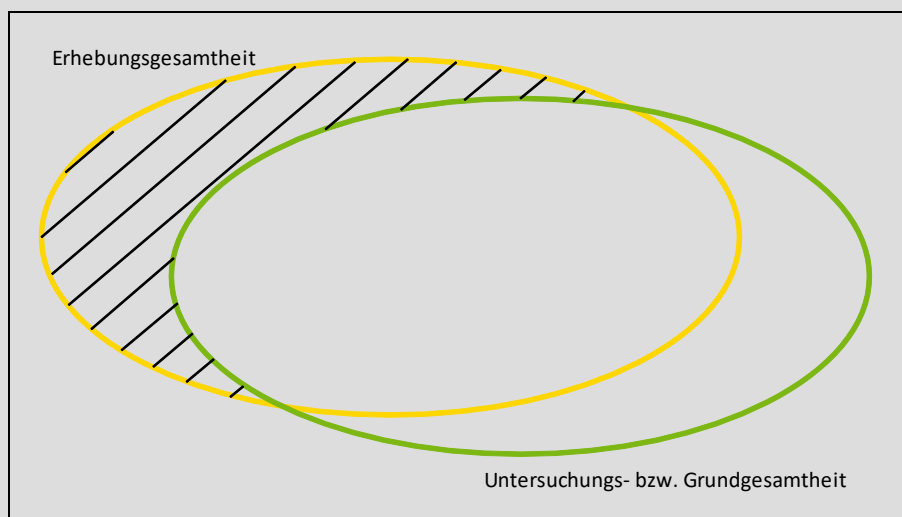
### 3.1.3 Defizit der Auswahlgrundlage: Existenz irrelevanter Hausumringe

Wie in Unterkapitel 3.1 erwähnt, besteht die Auswahlgrundlage aus 48.791.300 (aufbereiteten) Hausumringen. Von diesen ist vermutlich allerdings nur ein kleiner Teil in dem Sinne relevant, dass er Schnittflächen mit Nichtwohngebäuden aufweist. Wie groß der Anteil dieser und im Umkehrschluss der Anteil irrelevanter Hausumringe ist, ist unbekannt. Denn wie ebenfalls unter 3.1 ausgeführt, liegen für fast 20 Mio. der knapp 49 Mio. Hausumringe überhaupt keine bzw. keine ausreichend aussagekräftigen Nutzungsattribute (LoD1-Gebäudefunktionen) vor, die einen verhältnismäßig sicheren Rückschluss auf die Zugehörigkeit des zum jeweiligen Hausumring gehörenden Bauwerks zur Grundgesamtheit der

Nichtwohngebäude (geschweige denn zur besonders interessierenden Teilmenge der GEG-relevanten Nichtwohngebäude) zulassen. Und auch unter denjenigen Hausumringen, die sich anhand ihrer Nutzungsattribute als zu einem Nichtwohngebäude gehörig oder aber als irrelevant einstufen lassen, kann diese Einstufung in Einzelfällen falsch sein, insbesondere weil Nutzungsänderungen im Liegenschaftskataster erst zeitverzögert erfasst werden.<sup>15</sup>

In methodischer Terminologie ausgedrückt, ist somit davon auszugehen, dass der Auswahlgrundlage in substantiellem Umfang Erhebungseinheiten (d. h. Hausumringe) angehören, denen keine Untersuchungseinheiten (d. h. Nichtwohngebäude) zugeordnet sind, die für die vorliegende Untersuchung also irrelevant sind<sup>16</sup>. Die Auswahlgrundlage ist somit durch Übererfassung (*overcoverage*) gekennzeichnet. In Abbildung 3-9 ist das Phänomen der Übererfassung schematisch dargestellt. Die Menge irrelevanter Erhebungseinheiten entspricht dabei der schraffierten Fläche<sup>17</sup>.

**Abbildung 3-9: Schematische Darstellung von Übererfassung (schraffierter Bereich)**



Quelle: Eigene Darstellung

<sup>15</sup> Die derzeitigen technischen Möglichkeiten im Liegenschaftskataster lassen eine umgehende Erfassung von Nutzungsänderungen nicht zu.

<sup>16</sup> In einem Pfeildiagramm der Art von Abbildung 3-2 dargestellt, sind irrelevante Erhebungseinheiten solche, von denen kein auf eine Untersuchungseinheit weisender Pfeil abgeht.

<sup>17</sup> Abbildung 3-9 abstrahiert aus darstellerischen Gründen davon, dass im vorliegenden Fall Erhebungseinheiten (Hausumringe) und Untersuchungseinheiten (Nichtwohngebäude) Elemente unterschiedlicher Art sind und es somit streng genommen überhaupt keine Schnittmenge geben kann. Terminologisch richtig dürfte man in der Abbildung deshalb nicht von der Erhebungsgesamtheit sprechen. Denn gemeint sind nicht die die Erhebungsgesamtheit bildenden Hausumringe selbst, sondern vielmehr diejenigen Bauwerke, auf die diese Hausumringe zeigen, d. h. deren Grundrisse gemeinsame Schnittflächen mit Hausumringflächen aufweisen.

Übererfassung stellt kein gravierendes methodisches Problem dar, solange irrelevante Erhebungseinheiten als solche identifizierbar sind – wenn auch erst nach der Stichprobenziehung. Irrelevante Erhebungseinheiten gehören dann zwar zum Modellrahmen, wirken sich in methodischer Hinsicht jedoch kaum nachteilig auf diesen aus (vgl. Unterkapitel 3.3.1). Irrelevante Erhebungseinheiten sind vielmehr in erhebungspraktischer Hinsicht störend, weil das Erfordernis einer Relevanzfeststellung Erhebungsaufwand verursacht und weil bei gegebenem Umfang der Hausumring-Stichprobe jede als irrelevant aufgedeckte Erhebungseinheit, also jeder nicht mit Nichtwohngebäuden in Verbindung stehende Stichproben-Hausumring zulasten anderer relevanter Hausumringe und damit letztendlich zulasten der Zahl identifizierter Nichtwohngebäude geht. Anzustreben ist daher, die Hausumring-Stichprobe so zu konzipieren, dass die Zahl von Irrelevanz-Feststellungen gering gehalten wird. Gleichzeitig ist zu gewährleisten, dass über die gezogene Hausumring-Stichprobe möglichst viele GEG-relevante Nichtwohngebäude identifiziert werden können, weil dieses Nichtwohngebäudesegment aus inhaltlicher Sicht von besonderem Interesse ist. Beides ist jedoch nur möglich, wenn ex ante alle Hausumringe – auch die knapp 20 Mio. Hausumringe ohne ausreichend aussagekräftige Nutzungsattribute – im Hinblick auf ihre Zugehörigkeit zu GEG-relevanten Nichtwohngebäuden hin bewertet werden, und zwar möglichst metrisch in Gestalt von Wahrscheinlichkeiten. Während die Zuordnung solcher im Folgenden als „Relevanzwahrscheinlichkeiten“ bezeichneter Wahrscheinlichkeiten im Fall von Hausumringen auf der Hand liegt, die anhand ihrer Nutzungsattribute und bei Akzeptanz einer nicht ausschaltbaren Restunschärfe als in diesem Sinne relevant oder aber irrelevant anzusehen sind<sup>18</sup>, stellt sich für die erwähnten knapp 20 Mio. Hausumringe ohne aussagekräftige Nutzungsattribute die Frage nach der Vorgehensweise zur Zuweisung von Relevanzwahrscheinlichkeiten. Hierzu infrage kommen nämlich zwei Ansätze. Der eine Ansatz, den knapp 20 Mio. Hausumringe ohne ausreichend aussagekräftige Nutzungsattribute Relevanzwahrscheinlichkeiten zuzuordnen, besteht in einer heuristischen Einstufung. Vor allem aufgrund des immensen Arbeitsaufwandes der dazu erforderlichen Einzelbegutachtung, aber auch aufgrund der Schwierigkeit der Zuordnung metrischer Wahrscheinlichkeitswerte und der subjektiven Komponente wurde von dieser Option zugunsten einer anderen Vorgehensweise Abstand genommen.

Die andere und hier gewählte Vorgehensweise besteht darin, Relevanzwahrscheinlichkeiten im Zuge eines binär-logistischen Regressionsansatzes zuzuspielen. Hierzu wurde ein dreistufiger Ansatz gewählt, der die unterschiedlichen Informationsstände im Hinblick auf die erwähnten 20 Mio. Hausumringe, für die Relevanzwahrscheinlichkeiten zu schätzen sind, berücksichtigt und der als erklärende,

<sup>18</sup> Als (mit Blick auf die GEG-Relevanz und die Nichtwohngebäudeeigenschaft) relevant eingestufte Hausumringe bekommen eine Relevanzwahrscheinlichkeit von 1, als irrelevant eingestufte Hausumringe eine Relevanzwahrscheinlichkeit von 0 zugewiesen. Zu beachten ist dabei, dass auch solche Hausumringe eine Relevanzwahrscheinlichkeit von 0 erhalten, bei denen zwar die Nichtwohngebäudeeigenschaft als gegeben betrachtet werden kann, die GEG-Relevanz allerdings nach Lage der Dinge verneint werden muss. Wie die späteren Ausführungen zeigen, haben allerdings auch als irrelevant eingestufte Hausumringe eine (wenn auch im Regelfall geringe) positive Ziehungswahrscheinlichkeit.

d. h. exogene, Variablen auf Merkmale zurückgriff, die der Verbundpartner IÖR allen rund 49 Mio. Hausumringen zuspielte<sup>19</sup>.

Das angesprochene unterschiedliche Ausmaß an vorliegenden Informationen zur Einschätzung der Untersuchungsrelevanz geht aus Tabelle 3-2 hervor, in der alle 48.791.300 Hausumringe unter Rückgriff auf Nutzungsattribute aus dem Amtlichen Liegenschaftskatasterinformationssystem (ALKIS) bezüglich ihrer Untersuchungsrelevanz bewertet wurden. Demnach ist für 1.256.246 Hausumringe keinerlei Einstufung möglich; es ist noch nicht einmal gesichert, dass diese Hausumringe zu Gebäuden gehören, denn möglicherweise beziehen sie sich auf solche Bauwerke, bei denen es sich nicht um Gebäude handelt. Bei weiteren 4.127.620 Hausumringen kann man anhand der vorhandenen Nutzungsattribute zumindest davon ausgehen, dass sie Schnittflächen mit Gebäuden bilden, auch wenn offen bleibt, ob es sich um Wohn- oder Nichtwohngebäude handelt. Für weitere 13.739.590 Hausumringe kann sogar eine Zugehörigkeit zum Nichtwohngebäude-sektor als gesichert angesehen werden; die Frage nach der GEG-Relevanz kann anhand der verfügbaren Nutzungsattribute allerdings nicht beantwortet werden. Bei den restlichen 29.667.844 Hausumringen war – unter Inkaufnahme einer nie ganz ausschaltbaren Restunschärfe – dagegen auch eine Entscheidung dahingehend möglich: Konkret konnte bei 438.110 Hausumringen eine Zugehörigkeit zum Bestand an GEG-relevanten Nichtwohngebäuden konstatiert werden, während bei 29.229.734 aufgrund der vorgefundenen Nutzungsattribute eine GEG-Relevanz bzw. die Nichtwohngebäudeeigenschaft verneint wurde, weil sie z. B. zu GEG-irrelevanten Nichtwohngebäuden oder zu Wohngebäuden gehörten.

Die in Tabelle 3-2 dargestellte Aufspaltung dieser 29.229.734 als irrelevant eingestuften Hausumringe erfolgte mit Blick auf den dreistufigen Regressionsansatz, der angesichts unterschiedlicher Verteilungsmuster bundeslandspezifisch umgesetzt wurde. Zunächst wurde auf einer ersten Stufe eine Schätzgleichung ermittelt, um den 1.256.246 Hausumringen ohne jede Einstufungsmöglichkeit eine Wahrscheinlichkeit  $p_1$  dafür zuzuspielen, dass es sich bei den durch diese Hausumringe beschriebenen Bauwerken um Gebäude handelt – ganz gleich ob Wohn- oder Nichtwohngebäude. Zur Ermittlung dieser Schätzgleichung wurde in fast allen Bundesländern (zu den Ausnahmen siehe unten) auf alle anderen Hausumringe zurückgegriffen, wobei die endogene Dummy-Variable, d. h. die Zielvariable, für die Hausumringe ohne Gebäudezugehörigkeit die Ausprägung 0 und für alle anderen Hausumringe die Ausprägung 1 annimmt<sup>20</sup>.

<sup>19</sup> Bei diesen zugespielten Merkmalen handelt es sich zum einen um Kennziffern zur Beschreibung der geometrischen Eigenschaften der Hausumringe bzw. – falls vorhanden – seiner angrenzenden Nachbarn und zum anderen um semantische Flächeninformationen aus dem Amtlichen Topographisch-Kartographischen Informationssystem (ATKIS).

<sup>20</sup> Die Ausprägung „0“ lässt sich als Wahrscheinlichkeit interpretieren, nicht zu einem Gebäude zu gehören, während die Ausprägung „1“ die Wahrscheinlichkeit für das Komplementärereignis, also zu einem Gebäude gehörend, angibt.



**Tabelle 3-2: Wahrscheinlichkeiten, die den Hausumringen der Auswahlgrundlage über Nutzungsattribute (LoD1-Gebäudefunktionen) zugewiesen wurden**

| Beschreibung                                  | Anzahl     | Zugewiesene Wahrscheinlichkeit für die Zugehörigkeit zu einem ... |                                                                  |                                                                              |
|-----------------------------------------------|------------|-------------------------------------------------------------------|------------------------------------------------------------------|------------------------------------------------------------------------------|
|                                               |            | ... Gebäude                                                       | ... NWG unter der Bedingung, dass Gebäudeeigenschaft erfüllt ist | ... GEG-relevanten NWG unter der Bedingung, dass NWG-Eigenschaft erfüllt ist |
| Keine Einstufung möglich                      | 1.256.246  | unbekannt                                                         | unbekannt                                                        | unbekannt                                                                    |
| Zu Gebäude gehörend, aber unklar, ob NWG      | 4.127.620  | 1                                                                 | unbekannt                                                        | unbekannt                                                                    |
| Zu NWG gehörend, aber unklar, ob GEG-relevant | 13.739.590 | 1                                                                 | 1                                                                | unbekannt                                                                    |
| Zu GEG-relevantem NWG gehörend                | 438.110    | 1                                                                 | 1                                                                | 1                                                                            |
| Zu GEG-irrelevantem NWG gehörend*             | 7.621.500  | 1                                                                 | 1                                                                | 0                                                                            |
| Aus konzeptionellen Gründen gestrichen**      | 112.217    | 1                                                                 | 1                                                                | -                                                                            |
| Zu keinem Gebäude gehörend                    | 437.977    | 0                                                                 | -                                                                | -                                                                            |
| Zu Wohngebäude gehörend                       | 21.058.040 | 1                                                                 | 0                                                                | -                                                                            |

\* Incl. Garagen

\*\* Nur für die Pilotphasen-Bundesländer Nordrhein-Westfalen und Thüringen gebildet

Quelle: Eigene Darstellung

Ziel und Zweck der binär-logistischen Regressionsanalyse auf der zweiten Stufe bestand darin, unter der Bedingung des Vorliegens der Gebäudeeigenschaft eine Schätzgleichung zu ermitteln, um den Hausumringen ohne Einstufungsmöglichkeit (oberste Tabellenzeile) und denjenigen, bei denen lediglich die Zugehörigkeit zu einem Gebäude feststeht (zweitoberste Tabellenzeile), (bedingte) Wahrscheinlichkeiten für das Vorliegen der Nichtwohngebäudeeigenschaft zuzuspielen ( $p_2$ ). Die Schätzgleichung wurde – wiederum mit Ausnahme einiger Bundesländer (siehe unten) – unter Rückgriff auf diejenigen Hausumringe gebildet, die in der zweiten Tabellenspalte von rechts eine „0“ oder eine „1“ aufweisen, wobei diese Werte den Ausprägungen der endogenen Dummy-Variable entsprechen.

Die Schätzgleichung auf der dritten Stufe diente schlussendlich dazu, (bedingte) Wahrscheinlichkeiten für das Vorliegen von GEG-Relevanz unter der Voraussetzung der Nichtwohngebäudeeigenschaft zuweisen zu können. Die Schätzgleichung stützte sich für das Gros der Bundesländer (zu den Ausnahmen siehe unten) auf diejenigen Hausumringe, die als zu Nichtwohngebäuden gehörend identifiziert wurden und bei denen die Frage nach der GEG-Relevanz vorab entschieden, also bejaht oder verneint werden konnte. Als Ausprägung der endogenen Dummy-Variable erhielt die erstgenannte Gruppe eine 1, die zweitgenannte Gruppe eine 0.

Durch diese Schätzgleichung wurde allen Hausumringen der ersten drei Tabellenzeilen, also allen 19.123.456 Hausumringen ohne ausreichend aussagekräftige Nutzungsattribute, in Abhängigkeit ihres jeweiligen Sets an erklärenden Merkmalen die betreffende bedingte Wahrscheinlichkeit  $p_3$  zugeordnet.

Durch die beschriebene stufenweise binär-logistische Ermittlung von Relevanzwahrscheinlichkeiten lassen sich für die 19.123.456 Hausumringe ohne ausreichend aussagekräftige Nutzungsattribute Relevanzwahrscheinlichkeiten angeben. Diese sind in der rechten Spalte von Tabelle 3-3 formelmäßig ausgewiesen. Zu beachten ist, dass  $p_1$ ,  $p_2$  und  $p_3$  hausumringindividuell sind, wodurch die 19.123.456 Hausumringe der ersten drei Tabellenzeilen in Abhängigkeit ihrer jeweiligen exogenen Merkmale individuelle Relevanzwahrscheinlichkeiten aufweisen. Alle anderen Hausumringe haben in Abhängigkeit ihrer Nutzungsattribute stattdessen zugewiesene Relevanzwahrscheinlichkeiten von 0 oder 1.

**Tabelle 3-3: Zugewiesene bzw. zugeschätzte Relevanzwahrscheinlichkeiten**

| Beschreibung                                  | Zugewiesene oder zugeschätzte Wahrscheinlichkeit für die Zugehörigkeit zu einem ... |                                                                  |                                                                              | Relevanzwahrscheinlichkeit |
|-----------------------------------------------|-------------------------------------------------------------------------------------|------------------------------------------------------------------|------------------------------------------------------------------------------|----------------------------|
|                                               | ... Gebäude                                                                         | ... NWG unter der Bedingung, dass Gebäudeeigenschaft erfüllt ist | ... GEG-relevanten NWG unter der Bedingung, dass NWG-Eigenschaft erfüllt ist |                            |
| Keine Einstufung möglich                      | $p_1$                                                                               | $p_2$                                                            | $p_3$                                                                        | $p_1 p_2 p_3$              |
| Zu Gebäude gehörend, aber unklar, ob NWG      | 1                                                                                   | $p_2$                                                            | $p_3$                                                                        | $p_2 p_3$                  |
| Zu NWG gehörend, aber unklar, ob GEG-relevant | 1                                                                                   | 1                                                                | $p_3$                                                                        | $p_3$                      |
| Zu GEG-relevantem NWG gehörend                | 1                                                                                   | 1                                                                | 1                                                                            | 1                          |
| Zu GEG-irrelevantem NWG gehörend*             | 1                                                                                   | 1                                                                | 0                                                                            | 0                          |
| Aus konzeptionellen Gründen gestrichen**      | 1                                                                                   | 1                                                                | -                                                                            | 0                          |
| Zu keinem Gebäude gehörend                    | 0                                                                                   | -                                                                | -                                                                            | 0                          |
| Zu Wohngebäude gehörend                       | 1                                                                                   | 0                                                                | -                                                                            | 0                          |

\* Incl. Garagen

\*\* Nur für die Pilotphasen-Bundesländer Nordrhein-Westfalen und Thüringen gebildet

Quelle: Eigene Darstellung

Wie bereits erwähnt, wurden die via dreistufigem binär-logistischem Regressionsansatz ermittelten Relevanzwahrscheinlichkeiten im Regelfall dergestalt bundeslandspezifisch ermittelt, dass nur Hausumringe des betreffenden Bundeslandes in

die Regressionsschätzungen einfließen. Ausnahmen bildeten dabei Hessen und das Saarland, für die das länderspezifische Vorgehen zu keinen sinnvollen Relevanzwahrscheinlichkeiten geführt hat. Ursächlich hierfür ist in beiden Bundesländern innerhalb der als zu einem Nichtwohngebäude gehörend eingestuften Hausumringe das jeweilige Verhältnis zwischen solchen Hausumringen, die (mutmaßlich) zu einem GEG-relevanten Nichtwohngebäude gehören, zu solchen Hausumringen, auf die das nicht zutrifft, denn das jeweilige Verhältnis weicht erheblich vom bundesweiten Verhältnis ab. Während Letzteres bei rund 1:17 liegt<sup>21</sup>, beträgt es in Hessen ca. 1:963 und im Saarland etwa 4:1. Dieser Umstand hat zur Folge, dass auf der jeweiligen dritten Stufe der binär-logistischen Regressionsschätzung im Vergleich zu anderen Bundesländern in Hessen auffallend niedrige und im Saarland auffallend hohe (bedingte) Relevanzwahrscheinlichkeiten geschätzt wurden. In beiden Bundesländern wurde daher auf die Ermittlung der dreistufigen binär-logistischen Regressionsgleichung unter ausschließlicher Rückgriff auf bundeslandzugehörige Hausumringe verzichtet. Stattdessen wurde auf Regressionsgleichungen zurückgegriffen, die auf Basis der in der Pilotphase gescreeenten 4.000 Hausumringe<sup>22</sup> ermittelt wurden.

Abbildung 3-10 zeigt die Verteilung der binär-logistisch ermittelten Relevanzwahrscheinlichkeiten für die 19.123.456 Hausumringe mit fehlenden oder nur unzureichend aussagekräftigen Nutzungsattributen. Zu beachten ist dabei, dass es methodisch ausgeschlossen ist, dass binär-logistisch ermittelte Relevanzwahrscheinlichkeiten exakt 0 oder 1 betragen. Stattdessen sind sie stets positiv, aber kleiner als 1.

Die Verteilung der binär-logistisch geschätzten Relevanzwahrscheinlichkeiten ist stark rechtsschief, was ein Reflex der Eingangsdaten der Regressionsschätzung ist. Denn wie Tabelle 3-2 zu entnehmen ist, steht vergleichsweise wenigen Hausumringen, die anhand ihrer Nutzungsattribute als zu einem GEG-relevanten Nichtwohngebäude gehörend eingestuft wurden (438.110 Hausumringe), eine über große Zahl von als irrelevant bewerteten Hausumringen (29.229.734 Hausumringe) gegenüber. Die Ballung der geschätzten Relevanzwahrscheinlichkeiten am unteren Ende der Verteilung zeigt sich auch darin, dass die mediane Relevanzwahrscheinlichkeit mit 0,06 deutlich unter der mittleren (0,23) liegt.

Zur Vermeidung von Missverständnissen sei an dieser Stelle darauf hingewiesen, dass Relevanzwahrscheinlichkeiten nicht mit Ziehungswahrscheinlichkeiten verwechselt werden dürfen, weil Relevanzwahrscheinlichkeiten im Regelfall nicht mit Ziehungswahrscheinlichkeiten identisch sind. Relevanzwahrscheinlichkeiten dienen zum einen dazu, die Stichprobenziehung dahingehend zu gestalten, dass Hausumringe, die mit größerer (Relevanz-) Wahrscheinlichkeit zu einem GEG-relevanten Nichtwohngebäude gehören, eher in die Stichprobe aufgenommen werden als andere Hausumringe<sup>23</sup>. Wie dieses Ziel erreicht wurde, wird in

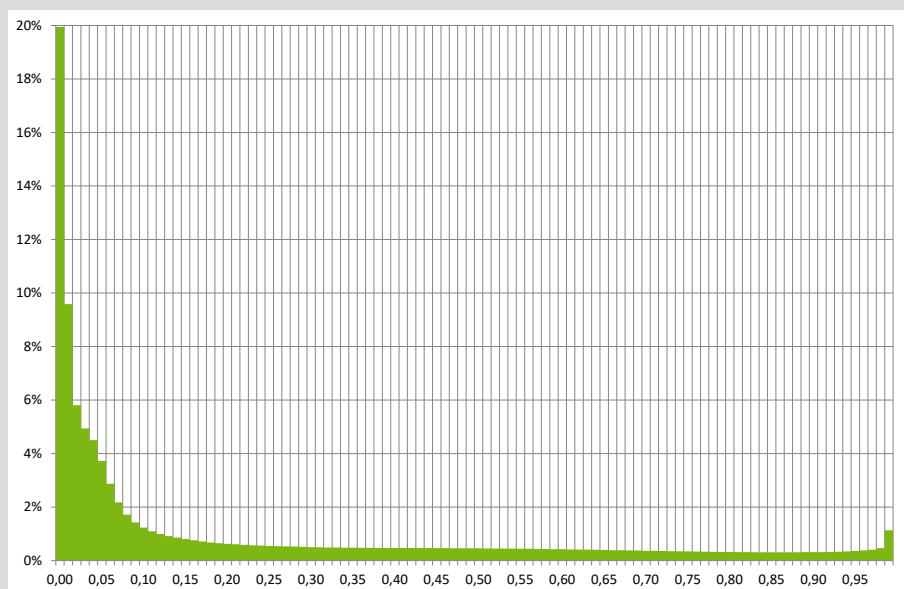
<sup>21</sup> Vgl. Tabelle 3-2: 438.110 Hausumringe, die als zu GEG-relevanten Nichtwohngebäuden gehörend eingestuft wurden, stehen 7.621.500 Hausumringen gegenüber, von denen auszugehen ist, dass sie zwar ebenfalls zu einem Nichtwohngebäude gehören, Letzteres aber nicht EV-relevant ist.

<sup>22</sup> Die 4.000 Hausumringe verteilen sich auf 20 Erhebungsbezirke in Nordrhein-Westfalen und Thüringen, die für die Pilotphase des Screenings aus den Erhebungsbezirksstichproben der beiden genannten Bundesländer ausgewählt wurden.

<sup>23</sup> Aufgrund dessen sind Relevanz- und Ziehungswahrscheinlichkeiten positiv miteinander korreliert.

Unterkapitel 3.2.3 erläutert. Zum anderen wurde zur Bildung von Erhebungsbezirken auf Relevanzwahrscheinlichkeiten zurückgegriffen (vgl. Unterkapitel 3.2.2).

**Abbildung 3-10: Verteilung der binär-logistisch ermittelten Relevanzwahrscheinlichkeiten**



Quelle: Eigene Berechnung und Darstellung

In diesem Kontext ist insbesondere zu beachten, dass auch den 29.229.734 Hausumringen mit zugewiesener Relevanzwahrscheinlichkeit von 0 positive Ziehungswahrscheinlichkeiten eingeräumt werden mussten. Denn zum einen ist die Einstufung als irrelevant mit Unsicherheit behaftet und kann daher in Einzelfällen falsch sein d. h. ein Hausumring mit einer zugewiesenen Relevanzwahrscheinlichkeit von 0 kann wider Erwarten doch zu einem GEG-relevanten Nichtwohngebäude gehören. Eine Aussonderung dieser Hausumringe (was äquivalent zu einer Ziehungswahrscheinlichkeit von 0 ist) würde daher der in Kapitel 2 aufgeführten Voraussetzung für erwartungstreue Schätzungen zuwider laufen, allen Nichtwohngebäuden (und damit auch allen GEG-relevanten Nichtwohngebäuden) eine Chance auf Aufnahme in die Stichprobe einzuräumen. Zum anderen besteht die Grundgesamtheit, über die am Ende datenorientierte Aussagen in Gestalt von populationsbeschreibenden Parametern getroffen werden sollen, nicht nur aus den besonders interessierenden GEG-relevanten Nichtwohngebäuden, sondern auch aus Nichtwohngebäuden, die GEG-irrelevant sind. Obwohl diesen – wie an späterer Stelle (vgl. Unterkapitel 3.2.3) noch erläutert wird – im Regelfall nur sehr geringe Ziehungswahrscheinlichkeiten eingeräumt werden, haben sie dennoch eine positive Chance auf Aufnahme in die Stichprobe. Dadurch kann auch für dieses

Gebäudesegment „Repräsentativität“ hergestellt werden<sup>24</sup>, auch wenn sich diesbezügliche Auswertungen auf vergleichsweise geringe Fallzahlen stützen müssen und die jeweiligen Standardfehler entsprechend hoch ausfallen werden.

Abschließend ist noch eine Entscheidung dahingehend zu treffen, wer beurteilt, ob ein Stichproben-Hausumring gemeinsame Flächen mit Nichtwohngebäuden aufweist und wenn ja, mit wie vielen und wie viele davon GEG-relevant sind, denn bei Letzteren sollen in der dem Screening nachgelagerten Breitenerhebungsphase durch eine Befragung des Gebäudeansprechpartners weitere Informationen über das Gebäude erhoben werden. Angesichts der ohnehin gebotenen Einrichtung eines Screenings lag es auf der Hand, zur Relevanzfeststellung der geschilderten Art das Screening-Personal einzubinden. Dabei sollte das Screening-Personal aber nicht eine finale Entscheidung hinsichtlich der Relevanz bzw. Irrelevanz treffen, sondern vielmehr vom öffentlichen Raum aus Angaben zur primären und sekundären Gebäudefunktion (jeweils differenziert nach Haupt- und Unterkategorie) machen<sup>25</sup>. Nach der zentralen Prüfung dieser Angaben durch das Projektteam des Verbundpartners BUW-ÖPB wurden im Bewusstsein, dass eine treffsichere Unterscheidung zwischen GEG-relevanten Nichtwohngebäuden und anderen Gebäuden vom öffentlichen Raum aus nicht in allen Fällen möglich ist<sup>26</sup>, sodann all diejenigen Gebäude an die Breitenerhebung übergeben, bei denen zumindest die Möglichkeit bestand, dass es sich um ein GEG-relevantes Nichtwohngebäude handelt<sup>27</sup>. Die finale Klärung erfolgte dann im Interview mit dem Gebäudeansprechpartner.

### 3.1.4 Defizit der Auswahlgrundlage: Untererfassung

Das Gegenstück zu Übererfassung ist Untererfassung (*undercoverage*), also der Umstand, dass nicht alle Untersuchungseinheiten bzw. Nichtwohngebäude über die Auswahlgrundlage erreichbar sind<sup>28</sup>. Abbildung 3-11 stellt Untererfassung schematisch dar: Die Menge der Untersuchungseinheiten, denen keine Erhebungseinheiten (also Hausumringe) gegenüberstehen, befinden sich im schraffierten Bereich.

Untererfassung ist prinzipiell ein schwerwiegendes methodisches Problem, da eine der in Kapitel 2 beschriebenen Voraussetzung für die Gewährleistung von erwartungstreuen Schätzungen sowie für die Quantifizierbarkeit der stichprobenbedingten Ergebnisunsicherheit, nämlich die Sicherstellung positiver Ziehungswahrscheinlichkeiten für alle Elemente der Grundgesamtheit, nicht erfüllt ist. Konkret führt Untererfassung stets dazu, dass sog. Merkmalssummen (*population totals*) systematisch unterschätzt werden, und zwar im vorliegenden Fall in Bezug auf all

<sup>24</sup> Zur Diskussion des Begriffes der „Repräsentativität“ vgl. Kapitel 2, insbesondere auch Fußnote 4.

<sup>25</sup> Die vorgegebenen Kategorien orientierten sich dabei am Bauordnungszuordnungskatalog.

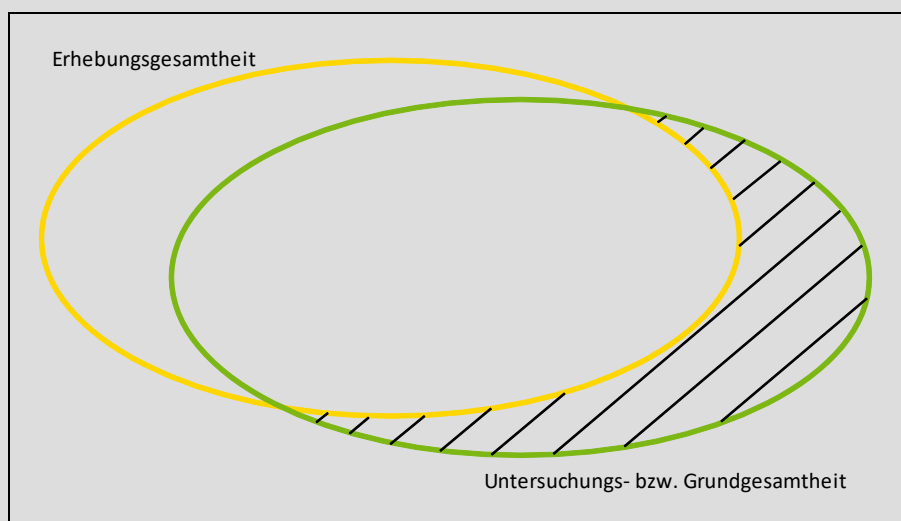
<sup>26</sup> Als Beispiel seien Gebäude angeführt, die teils Wohn-, teils gewerblichen Zwecken dienen. Eine Zugehörigkeit zum Nichtwohngebäudesektor ist nach der auch hier zugrunde gelegten Definition der amtlichen Statistik nur dann gegeben, wenn mehr als die Hälfte der Gesamtnutzfläche für Nichtwohnzwecke bestimmt ist.

<sup>27</sup> Nur Gebäude, bei denen nach dem Screening mit hinreichender Sicherheit die Frage nach der GEG-Relevanz und/oder der Nichtwohngebäudeeigenschaft verneint werden konnte, wurden vorab aussortiert.

<sup>28</sup> In einem Pfeildiagramm der Art von Abbildung 3-2 dargestellt, wären solche Untersuchungseinheiten dadurch gekennzeichnet, das auf sie kein von einer Erhebungseinheit abgehender Pfeil weist.

diejenigen Typen von Nichtwohngebäuden, in denen es Untererfassungen gibt. Da zu den Merkmalssummenschätzungen auch die Schätzung der Anzahl aller Nichtwohngebäude bestimmter Typen und darüber die Schätzung der Gesamtzahl aller Nichtwohngebäude gehört<sup>29</sup>, führt Untererfassung zu einer systematischen Unterschätzung des Bestands an Nichtwohngebäuden (und damit auch an GEG-relevanten Nichtwohngebäuden). Das Ausmaß der systematischen Unterschätzung von Merkmalssummen ist davon abhängig, wie stark die Auswahlgrundlage von Untererfassung betroffen ist.

**Abbildung 3-11: Schematische Darstellung von Untererfassung (schraffierter Bereich)**



Quelle: Eigene Darstellung

Neben der systematischen Unterschätzung von Merkmalssummen kann Untererfassung auch zu systematisch verzerrten Verhältnisschätzungen (*ratio estimations*) führen, insbesondere zu systematisch verzerrten Schätzungen zum Anteil von Nichtwohngebäuden mit einer bestimmten Eigenschaft<sup>30</sup>. Hierzu kommt es dann, wenn sich diejenigen Nichtwohngebäude, die über die Auswahlgrundlage nicht erreichbar sind, im Hinblick auf interessierende Auswertungsmerkmale systematisch von den anderen Nichtwohngebäuden unterscheiden. Das jeweilige Ausmaß der systematischen Verzerrung bei Verhältnisschätzungen hängt wiederum vom quantitativen Ausmaß der Untererfassung in der Auswahlgrundlage ab,

<sup>29</sup> Dass die Schätzung der Anzahl der Nichtwohngebäude bestimmter Typen eine Merkmalssummenschätzung darstellt, ist damit zu begründen, dass im Rahmen der Schätzung alle typenzugehörigen Nichtwohngebäude gedanklich mit 1 („trifft zu“) und der Rest mit 0 („trifft nicht zu“) gewertet und im Anschluss daran gewichtet aufsummiert werden, wobei als Gewichte die Kehrwerte der jeweiligen gebäudespezifischen Ziehungswahrscheinlichkeiten dienen (vgl. Unterkapitel 3.3.1).

<sup>30</sup> In Abhängigkeit vom auszuwertenden Merkmal kann sich die systematisch verzerrte Schätzung in einer systematischen Über- oder Unterschätzung äußern.

aber auch davon, wie stark sich die über die Auswahlgrundlage nicht erreichbaren Nichtwohngebäude von den erreichbaren abheben.

Auch wenn es sich bei der bisweilen erhobenen Forderung, wonach eine Auswahlgrundlage wegen der beschriebenen negativen Konsequenzen überhaupt keine Untererfassung aufweisen dürfe, um eine Maximalforderung handelt, die in der Erhebungspraxis nur in den seltensten Fällen vollumfänglich erfüllt ist, ist es gleichwohl unabdingbar, Untererfassung weitestgehend zurückzudrängen.

Scheidet – wie im vorliegenden Fall – die Möglichkeit aus, die Auswahlgrundlage durch eigene Nacherhebungen zu ergänzen und damit eine etwaige Untererfassung zu beseitigen, bleibt nur übrig, sich der „Lücken“ der Auswahlgrundlage bewusst zu sein und die Definition der Grundgesamtheit an den Möglichkeiten und Grenzen der Auswahlgrundlage auszurichten. Denn die quantitative Bedeutung von Untererfassung und die damit einhergehende Problemschwere bei Merkmalssummen- und ggf. auch Verhältnisschätzungen lassen sich über die Definition der Grundgesamtheit gezielt steuern. Dazu bietet es sich an, sich klar zu machen, welche Untersuchungseinheiten, hier Nichtwohngebäude, über die Auswahlgrundlage überhaupt erreichbar sind. Erreichbar sind im vorliegenden Fall alle zum Zeitpunkt des Screenings<sup>31</sup> vorhandenen Nichtwohngebäude, die Schnittflächen mit Hausumringen haben, die im April 2015 in der vom Verbundpartner IÖR aufbereiteten Datenbank amtlicher Hausumringe Deutschland enthalten waren. Erreichbar sind somit auch nach dem Stichtag des Datenbankauszugs fertiggestellte Nichtwohngebäude, aber nur, solange sie Schnittflächen mit Hausumringen haben, die bereits zum Stichtag in der Datenbank vorhanden waren. Bei diesen über die Auswahlgrundlage erreichbaren Nichtwohngebäuden mit Fertigstellungszeitpunkt nach April 2015 dürfte es sich deshalb fast ausschließlich um solche Neubauten handeln, die auf Abrissgrundstücken erstellt wurden (vgl. auch Fußnote 13). Neubauten auf neu ausgewiesenen Bauflächen haben dagegen grundsätzlich keine Erfassungschance. Gleiches gilt auch für Nichtwohngebäude, die im April 2015 zwar bereits errichtet waren, die aber aufgrund der zeitverzögerten, sich auf bis zu drei Jahre hinziehenden Aufnahme in das Liegenschaftskataster dort im April 2015 noch nicht erfasst waren. Auch was Nichtwohngebäude mit sehr kleinen Grundflächen angeht, kann nicht davon ausgegangen werden, dass diese allesamt durch die Auswahlgrundlage „abdeckt“ sind, also über Hausumringe erreichbar sind. Beispielsweise sind dem Hessischen Vermessungs- und Geoinformationsgesetz (HVGG) zufolge im Liegenschaftskataster Liegenschaften für das gesamte Landesgebiet zwar flächendeckend und vollständig nachgewiesen; allerdings zählen zu diesen Liegenschaften neben allen Flurstücken nur solche Gebäude, die „liegenschaftsrechtlich bedeutsam“ sind (vgl. § 9 Abs. 1 HVGG). Dabei handelt es sich entsprechend der Ausführungsverordnung zum erwähnten Gesetz um Gebäude mit einer Mindestfläche von 20 m<sup>2</sup> (falls innerhalb im Zusammenhang bebauter Ortsteile oder rechtskräftiger Bebauungspläne gelegen) bzw. 50 m<sup>2</sup> (andernfalls). Zwar können gemäß § 9 Abs. 8 HVGG auch Gebäude unterhalb dieser Flächennormen im Liegenschaftskataster geführt werden, aber nur, wenn der katasterführenden Behörde geeignete Unterlagen zur Verfügung gestellt werden. Im Ergebnis kann daher nicht von einer vollständigen Erfassung ausgegangen werden.

---

<sup>31</sup> Zu beachten ist, dass sich das Screening aus erhebungspraktischen Gründen über einen längeren Zeitraum bis Ende Mai 2019 erstreckte.



Zusammenfassend dürften Nichtwohngebäude, die über die Auswahlgrundlage nicht erreicht werden können, somit vor allem Gebäude mit jüngerem Fertigstellungsdatum bzw. mit sehr kleinen Grundflächen sein.

Die geschilderten Zusammenhänge legen somit nahe, als Grundgesamtheit die Gesamtheit aller Nichtwohngebäude zu definieren, die über die Auswahlgrundlage erreichbar sind. Diese Festlegung hat den Vorteil, dass es notwendigerweise keinerlei Untererfassung geben kann, bringt aber den Nachteil mit sich, dass sie inhaltlich unbefriedigend erscheinen mag. Jede rein inhaltlich motivierte Festlegung der Grundgesamtheit, beispielsweise als die Gesamtheit aller Nichtwohngebäude mit Baujahr bis einschließlich 2014 oder als die Gesamtheit aller gegenwärtig vorhandener Nichtwohngebäude, wäre aber auch mit nachteiligen Konsequenzen verbunden gewesen, und zwar in jedem Fall mit dem Eingeständnis von Untererfassung. Und wenn wie im ersten Beispiel eine Baualtersgrenze gezogen worden wäre, hätte dies folgerichtig darauf hinauslaufen müssen, nach dieser Baualtersgrenze erstellte Nichtwohngebäude im Screening als solche zu kennzeichnen und im Zuge der nachgelagerten Datenaufbereitung aus der Untersuchung auszuschließen. Dies wäre allerdings dem Bestreben zuwidergelaufen, in spätere Auswertungen so viele Nichtwohngebäude wie möglich einzubinden.

Die hier angelegte Definition der Grundgesamtheit lässt somit auch Auswertungen zu, die jüngere Gebäudebaualter (also Baujahr nach 2014) oder Nichtwohngebäude mit sehr kleinen Grundflächen einschließen. In dem Maße, in dem sich Auswertungen auf solche speziellen Segmente des Nichtwohngebäudebestands beschränken, steigt allerdings das Risiko, dass die tatsächlichen Gegebenheiten in den betreffenden Segmenten verzerrt wiedergegeben werden.

### **3.1.5 Defizit der Auswahlgrundlage: Fehlende Gebäudemerkmale**

Eine ideale Auswahlgrundlage würde im vorliegenden Fall ein Set von grundlegenden Gebäudemerkmalen vorhalten, so dass direkt aus der Datenbank Strukturinformationen über die Grundgesamtheit abgeleitet werden können. Damit einhergehen würden drei Vorteile. Erstens bestünde die Möglichkeit, ohne die Notwendigkeit einer Stichprobenziehung und daher ohne die damit verbundene Unsicherheit die Grundgesamtheit entlang dieser grundlegenden Gebäudemerkmale zu beschreiben. Zweitens könnte unter Rückgriff auf diese Informationen die Schätzgenauigkeit derjenigen Parameter erhöht werden, zu denen die Datenbank keine Informationen bereithält und die daher nur auf Stichprobenbasis bestimmt werden können. Hierzu könnte bereits die Ziehung der Stichprobe beitragen, indem entsprechend geschichtet gezogen wird. Und auch bei der sich anschließenden Parameterschätzung könnten diese Strukturinformationen „gewinnbringend“ eingesetzt werden. Und drittens würden diese Strukturinformationen erlauben, etwaige spätere Totalausfälle von Untersuchungsgebäuden wegen sog. Nichtbeobachtbarkeit (*unit nonresponse*), also beispielsweise aufgrund der Verweigerung einer Interviewteilnahme durch den Gebäudeansprechpartner, dahingehend zu analysieren, ob diese Ausfälle systematischer Natur sind. Im Bedarfsfall werden dadurch die Möglichkeiten geschaffen gegenzusteuern, und zwar sowohl während der Feldphase durch geeignete erhebungspraktische Maßnahmen als auch ex post durch entsprechende stichprobenmethodische Werkzeuge (vgl. Unterkapitel 4.1.2).



Da es mit vertretbarem Aufwand nicht möglich war, jeden Hausumring der Datenbank um solche Gebäudemerkmale anzureichern, wurde das ohnehin notwendige Screening dazu genutzt, um für die dort identifizierten Gebäude einige grundlegende und vom öffentlichen Raum aus erkennbare Merkmale mit zu erheben. Denn dadurch wird die Möglichkeit geschaffen, den gerade beschriebenen Ausfall von Untersuchungsgebäuden in der Breitenerhebung zu analysieren und ggf. zu korrigieren. Zu diesem methodischen Zusatznutzen tritt ein rein inhaltlich motivierter hinzu. Denn durch das Erheben grundlegender Gebäudemerkmale liefert das Screening auf – wie sich noch herausstellen wird – sehr breiter Stichprobenbasis Daten, dank derer wesentliche Strukturen der Grundgesamtheit mit verhältnismäßig geringer Stichprobenbedingter Unsicherheit verzerrungsfrei beschrieben werden können.

In der Gesamtbetrachtung gelingt es durch das Screening, Defizite der herangezogenen Auswahlgrundlage zu beseitigen, zurückzudrängen oder wenigstens dadurch handhabbar zu machen, dass nicht zu beseitigende Defizite aufgedeckt und so einer Problemlösung zugänglich gemacht werden<sup>32</sup>. Wie sich diese Problemlösung im Einzelnen darstellt, ist Gegenstand nachfolgender Ausführungen. Zunächst aber wird das Stichprobenauswahlverfahren erläutert. Dabei wird sich zeigen, dass sich die Einrichtung einer Screeningphase nachhaltig auf das Ziehungsverfahren ausgewirkt hat.

## 3.2 Stichprobenauswahlverfahren

Die Festlegung des Stichprobenauswahlverfahrens erfordert insbesondere Entscheidungen, wie viele Erhebungseinheiten, im vorliegenden Fall Hausumringe, zu ziehen sind (vgl. Unterkapitel 3.2.1) und wie die Ziehungen zu gestalten sind (vgl. Unterkapitel 3.2.2 und 3.2.3). Durch das Stichprobenauswahlverfahren in Verbindung mit dem jeweils vorliegenden Hausumring-Gebäude-Beziehungsgeflecht ergeben sich letztlich die zur Schätzung von populationsbeschreibenden Parametern einschließlich der dazugehörigen Standardfehler erforderlichen Gebäudeziehungswahrscheinlichkeiten (vgl. Unterkapitel 3.2.4).

### 3.2.1 Stichprobenumfang

Auch wenn sich die Stichprobenziehung auf Hausumringe als Erhebungseinheiten erstreckt, richtet sich das inhaltliche Erkenntnisinteresse auf Nichtwohngebäude als Untersuchungseinheiten der vorliegenden Erhebung. Ein besonderer Fokus ist dabei auf GEG-relevante Nichtwohngebäude gerichtet. Aufgrund dessen ist für die Festlegung des Hausumring-Stichprobenumfangs maßgebend, auf wie viele Datensätze GEG-relevanter Nichtwohngebäude sich die späteren Auswertungen zum Modernisierungsgeschehen idealerweise stützen sollen, d. h. wie groß der Nettostichprobenumfang in Bezug auf GEG-relevante Nichtwohngebäude im Idealfall sein soll.

Wie bereits am Ende von Kapitel 2 ausgeführt, steht die Festlegung des Stichprobenumfangs im Spannungsfeld zweier konkurrierender Ziele, nämlich zum einen

---

<sup>32</sup> So erlaubt das Screening beispielsweise, das konkrete Beziehungsgeflecht zwischen dem gezogenen Hausumring und dem bzw. den mit diesem Hausumring verbundenen Nichtwohngebäude(n) aufzudecken (vgl. Unterkapitel 3.1.2).

dem Ziel, die stichprobenbedingte Ergebnisunsicherheit zu minimieren, und zum anderen dem Ziel, die Erhebungskosten möglichst gering zu halten.

Dieser Zielkonflikt wurde durch die ambitionierte Vorgabe eines anzustrebenden Nettostichprobenumfangs von möglichst 10.000 GEG-relevanten Nichtwohngebäuden aufgelöst. Denn dadurch können die typischerweise im unteren einstelligen Prozentbereich liegenden, bauteilbezogenen Modernisierungsquoten auch für Teilssegmente des Gebäudebestands mit vertretbarer statistischer Unsicherheit bestimmt werden, sofern die Untergruppen ausreichend viele GEG-relevante Nichtwohngebäude umfassen. Die angestrebte Zahl von 10.000 Gebäudedatensätzen mit validen Informationen zum Modernisierungsgeschehen, aber auch zu typologischen Merkmalen, zum Wärmeschutz, zur Gebäudetechnik sowie zum Bewirtschaftungsverhalten und den Eigentümerstrukturen ist hinsichtlich der Aussagekraft der Ergebnisse, aber auch hinsichtlich der damit verbundenen Erhebungskosten eine sinnvoll austarierte Größe. Und selbst wenn diese ambitionierte Größenordnung – wie sich im Nachhinein herausstellte – doch nicht erreicht werden kann und man sich mit weniger Gebäudedatensätzen begnügen muss, ist damit noch nicht das Untersuchungsziel gefährdet.

Was die Erhebung von Energieverbrauch und -bedarf angeht, wäre ein ebenso großer Nettostichprobenumfang fraglos wünschenswert. Angesichts des ungleich höheren Erhebungsaufwandes, der erforderlichen Fachkompetenz des Interviewers und nicht zuletzt mit Blick auf die Auskunftsbereitschaft der Befragten fiel die Entscheidung, erhebungsorganisatorisch zwischen einer Breiten- und einer nachgelagerten Tieferhebung zu unterscheiden und die Erhebung von Energieverbrauch und -bedarf auf die Tieferhebung von bis zu 1.000 Gebäuden zu beschränken, die sich aus den „Breitenerhebungsgebäuden“ rekrutieren.

Um die mehrheitlich nur auf dem Befragungswege zu erhebenden Daten für im Idealfall 10.000 GEG-relevante Nichtwohngebäude zu erreichen, ist eine wesentlich größere Bruttostichprobe erforderlich, denn erfahrungsgemäß kommt aus unterschiedlichen Gründen nur in einem kleinen Teil der Fälle ein auswertbares Interview mit einem auskunftswilligen Gebäudeansprechpartner zustande. Zur Beantwortung der Frage, wie groß der entsprechende Hebel sein muss, existieren mangels vergleichbarer Vorgängeruntersuchungen jedoch keine belastbaren Erfahrungen, so dass ersatzweise auf Erfahrungswerte aus – thematisch allerdings anders gelagerten – schriftlich-postalischen Gebäude- bzw. Wohnungseigentümergebefragungen zurückgegriffen wurde, wo gewöhnlich von Response- bzw. Rücklaufquoten in einer Größenordnung von rund 20% berichtet wird<sup>33</sup>. Dies führte zu dem Ziel, im der Breitenerhebung vorgelagerten Screening möglichst 50.000 mutmaßlich GEG-relevante Nichtwohngebäude zu identifizieren<sup>34</sup>, um für

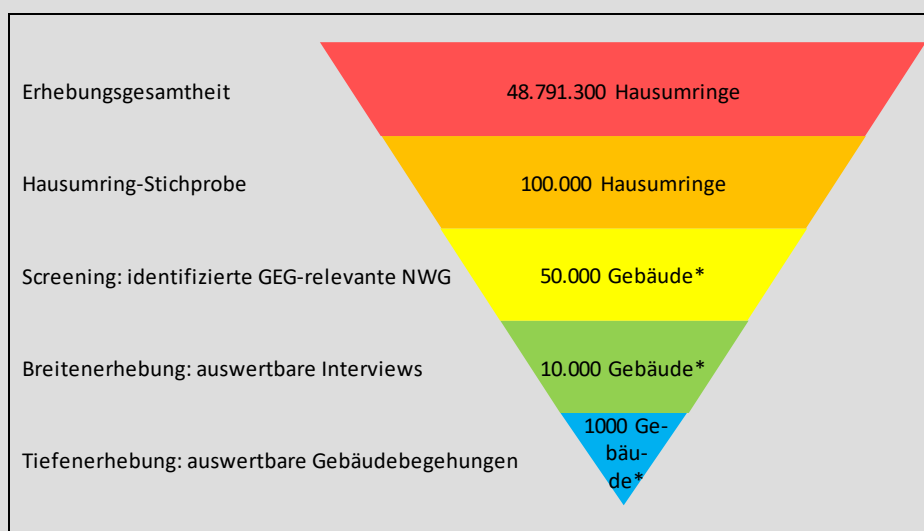
<sup>33</sup> So berichten (Cischinsky et al. 2015 S. 37) im Rahmen einer bundesweiten schriftlich-postalischen Befragung der Eigentümer von Mietwohnungen in Mehrfamilienhäusern von einer Rücklaufquote (bezogen auf den auswertbaren Rücklauf) von 22,0%. (Cischinsky, Diefenbach 2018 S. 34) erreichten bei einer bundesweiten Wohnungsbäudeerhebung, bei der ebenfalls Eigentümer schriftlich-postalisch befragt wurden, eine Rücklaufquote von 18,2%.

<sup>34</sup> Eine treffsichere Feststellung der GEG-Relevanz ist vom öffentlichen Raum aus nicht möglich, so dass im Screening nur entsprechende Mutmaßungen angestellt werden konnten. Eine endgültige Entscheidung hinsichtlich der GEG-Relevanz konnte daher erst im Zuge der Befragung während der Breitenerhebung getroffen werden. Vergleichbares galt auch für die Abgrenzung zwischen Wohn- und Nichtwohngebäuden im Falle gemischt genutzter Gebäude, die im Einklang mit der Definition der amtlichen Statistik nur anhand der Gegenüberstellung der

die sich anschließende Breitenerhebung einen ausreichend großen Vorrat an Untersuchungsgebäuden vorhalten zu können. Da sich die Stichprobenziehung auf Hausumringe als Erhebungseinheiten erstreckte, war eine Entscheidung dahingehend zu treffen, wie viele Hausumringe zu ziehen sind, um diese Zielvorgabe zu erfüllen. Zum einen aufgrund der nicht notwendigerweise eindeutigen Beziehung zwischen Hausumring und Nichtwohngebäude und zum anderen aufgrund der methodisch gebotenen Notwendigkeit, auch Hausumringe mit Relevanzwahrscheinlichkeiten kleiner als 1 in die Hausumring-Stichprobe aufzunehmen<sup>35</sup>, fiel die Entscheidung, doppelt so viele Hausumringe zu screenen wie mutmaßlich GEG-relevante Nichtwohngebäude identifiziert werden sollen, d. h. für das Screening eine Stichprobe von 100.000 Hausumringen zu ziehen.

Abbildung 3-12 veranschaulicht das Mengengerüst der gezogenen Hausumring-Stichprobe und der daraus angestrebten Zahl von (auswertbaren) Gebäudedaten-sätzen.

**Abbildung 3-12: Hausumring-Stichprobe und angestrebte Zahl von Gebäudedatensätzen**



\*angestrebt

Quelle: Eigene Darstellung

Wie ausgeführt, wurde der Hausumring-Stichprobenumfang auf Basis von Mutmaßungen insbesondere zur Teilnahmebereitschaft kontaktierter

entsprechenden Nutzflächen vorgenommen werden kann (vgl. auch Fußnote 26). Aus methodischen Gründen wurden alle im Screening identifizierten Gebäude, bei denen zumindest die Möglichkeit bestand, dass es sich um GEG-relevante Nichtwohngebäude handelt, in die Bruttostichprobe der Breitenerhebung aufgenommen (sog. breitenerhebungsrelevante Gebäude, vgl. Abbildung 4-1), um dort eine endgültige Klärung herbeizuführen.

<sup>35</sup> Anderweitig wäre die in Kapitel 2 aufgelistete dritte Voraussetzung zur Herstellung von Erwartungstreue und zur Schätzung von Standardfehlern, nämlich die Gewährleistung positiver Ziehungswahrscheinlichkeiten für alle Nichtwohngebäude der Grundgesamtheit, verletzt worden.

Gebäudeeigentümer und zu typischen Beziehungsgeflechten zwischen Hausumringen und Gebäuden abgeleitet, da belastbare empirische Vorerfahrungen nicht vorlagen. Dass sich diese Mutmaßungen im Nachhinein nur näherungsweise erfüllten und die angestrebten Fallzahlen am Ende nicht erreicht werden konnten (vgl. auch Abbildung 4-1), stellt die beschriebene Vorgehensweise zur Ableitung des Hausumring-Stichprobenumfangs jedoch nicht infrage. Zu beachten ist auch, dass die ambitionierten Zielvorgaben mit ausreichend Puffer ausgestattet waren, so dass eine Zielverfehlung nicht gleich das gesamte Forschungsvorhaben gefährdet. Und tatsächlich sind die letztlich erzielten Fallzahlen immer noch groß genug, um die gewünschten Ergebnisauswertungen vornehmen zu können.

### 3.2.2 1. Ziehungsstufe: Geschichtete Zufallsauswahl von Erhebungsbezirken

Wie in Unterkapitel 3.1 dargelegt, erfordern diverse Defizite der verwendeten Auswahlgrundlage ein Vor-Ort-Aufsuchen einer ausreichend großen Hausumring-Stichprobe. Dieser Umstand legt mit Rücksicht auf die begrenzte räumliche Mobilität des hierfür eingesetzten Personals sowie zum Zwecke der Minimierung der Wegekosten nahe, der Ziehung von Hausumringen eine Ziehung von flächenmäßig begrenzten Erhebungsbezirken voranzustellen. Dadurch entsteht ein zweistufiges Ziehungsverfahren, bei dem auf der ersten Ziehungsstufe zunächst deutschlandweit eine ausreichend große Zahl von Erhebungsbezirken und erst auf der zweiten Ziehungsstufe Hausumringe innerhalb dieser Erhebungsbezirke zufällig ausgewählt werden. Mit der Vorschaltung von Erhebungsbezirken tritt eine weitere Art von Erhebungseinheiten hinzu, die – weil auf der ersten Ziehungsstufe angesiedelt – in der Stichprobentheorie als sog. primäre Erhebungseinheiten (*primary sampling units*, abgekürzt *PSUs*) bezeichnet werden. Folgerichtig fungieren Hausumringe im vorliegenden Fall als sog. sekundäre Erhebungseinheiten (*secondary sampling units*, abgekürzt *SSUs*), da sie die zu ziehenden Einheiten auf der nachgelagerten zweiten Ziehungsstufe darstellen<sup>36</sup>.

#### Bildung von Erhebungsbezirken

Die vorgeschaltete Ziehung von Erhebungsbezirken erfordert drei Festlegungen. Die erste Festlegung bezieht sich darauf, was als Erhebungsbezirk herangezogen wird, insbesondere ob vorhandene räumliche Aggregate wie beispielsweise Städte und Gemeinden sich als Erhebungsbezirke eignen. Die zweite Festlegung betrifft die Anzahl der auf der ersten Stufe zu ziehenden Erhebungsbezirke und die dritte das konkrete Ziehungsverfahren.

Was die Festlegung von Erhebungsbezirken angeht, ist zwischen einer Muss- und einer Sollvorgabe zu unterscheiden. Aus methodischen Gründen unbedingt einzuhalten ist die Vorgabe, dass jeder Hausumring einem, und zwar nur einem einzigen Erhebungsbezirk zugeordnet sein darf. Mengentheoretisch ausgedrückt, müssen die Erhebungsbezirke somit eine Zerlegung der Erhebungsgesamtheit aller Hausumringe bilden, was in erhebungspraktischer Hinsicht darauf hinausläuft, die Fläche Deutschlands vollständig und überlappungsfrei in Erhebungsbezirke aufzuteilen.

<sup>36</sup> Da im vorliegenden Fall die zweite Ziehungsstufe gleichzeitig auch die unterste Ziehungsstufe ist, können die Hausumringe auch als *ultimate sampling units* bezeichnet werden (Särndal et al. 1992 S. 125).

Die Sollvorgabe besteht im vorliegenden Fall dagegen in der Forderung, dass die Erhebungsbezirke nicht unnötig, aber dennoch ausreichend „groß“ sein sollten. „Groß“ hat im vorliegenden Kontext zwei verschiedene Bedeutungen. „Nicht unnötig groß“ ist in Bezug auf die Fläche zu verstehen, denn flächenmäßig zu große Erhebungsbezirke stehen dem Ziel der Wegekostenminimierung entgegen und stellen letztlich die mit der Vorschaltung von Erhebungsbezirken verbundene Intention infrage. Die Erhebungsbezirke sollten daher flächenmäßig so klein wie möglich sein. Die Forderung nach „ausreichend großen“ Erhebungsbezirken bezieht sich dagegen auf die Zahl der darin enthaltenen GEG-relevanten Nichtwohngebäude. Denn in diesem Sinne zu kleine Erhebungsbezirke führen dazu, dass zur Erreichung des Ziels der Identifizierung von mindestens 50.000 mutmaßlich GEG-relevanten Nichtwohngebäuden eine umso größere Zahl von Erhebungsbezirken zu ziehen ist, was den organisatorischen Aufwand der Screenerrekrutierung, -schulung und -überwachung erhöht. Die Forderung nach „ausreichend großen“ Erhebungsbezirken impliziert auch eine gewisse Homogenität der Erhebungsbezirke in Bezug auf die eingebundene Zahl GEG-relevanter Nichtwohngebäude. Letztere sollte daher nicht unnötig stark zwischen den Erhebungsbezirken fluktuieren, denn dann ist es möglich, die Erhebungsbezirke ziehungstechnisch dahingehend gleich zu behandeln, dass für jeden Erhebungsbezirk derselbe Hausumring-Stichprobenumfang vorgesehen werden kann<sup>37</sup>. Damit konfrontiert, eignen sich die als Beispiel angeführten Städte- und Gemeindegebiete aufgrund ihrer stark unterschiedlichen Bestände an GEG-relevanten Nichtwohngebäuden kaum als Erhebungsbezirke: Will man im Hinblick auf den Bestand an GEG-relevanten Nichtwohngebäuden vergleichbar „große“ Erhebungsbezirke, sind Großstädte in diesem Sinne zu groß und Gemeinden oftmals zu klein. Stattdessen sollten die Erhebungsbezirke in gewerblich genutzten Gebieten relativ klein ausfallen, während in ländlichen Gebieten ohne intensive gewerbliche Nutzung größere, bisweilen mehrere Gemeinden umfassende Erhebungsbezirke vorzusehen sind – gleichwohl unter Beachtung der Maßgabe, dass zu große Erhebungsbezirke dem Wegekostenminimierungsziel zuwider laufen. In diesem Sinne wurden die Erhebungsbezirke eigens für die Belange der vorliegenden Untersuchung gebildet, wozu bis auf die Wahrung der Bundeslandgrenzen administrative Grenzen nicht weiter beachtet wurden. Die Beachtung der Bundeslandgrenzen war damit begründet, dass die binär-logistische Schätzung von Relevanzwahrscheinlichkeiten, auf die zur weiter unten beschriebenen Erhebungsbezirksbildung zurückgegriffen wurde, ebenfalls bundeslandspezifisch erfolgte und es von Bundesland zu Bundesland teils deutliche Unterschiede im Verteilungsmuster der Relevanzwahrscheinlichkeiten gibt (vgl. Unterkapitel 3.1.3).

Was in Bezug auf den Bestand an GEG-relevanten Nichtwohngebäuden „ausreichend“ (und vergleichbar) groß ist, bedarf noch einer Quantifizierung. Unter der Vorgabe einer 100.000er Hausumring-Stichprobe, anhand derer im Screening

<sup>37</sup> Methodischer Hintergrund ist das Bestreben, unter sonst gleichen Bedingungen jedem GEG-relevanten Nichtwohngebäude in Deutschland in etwa dieselbe Ziehungswahrscheinlichkeit zuordnen zu können. Das ließe sich zwar auch bei starken Unterschieden in der Erhebungsbezirksgröße realisieren, hätte dann aber zur Folge, dass die erhebungsbezirksspezifischen Hausumring-Stichprobenumfänge individuell anzupassen wären. Aufgrund der Unkenntnis, welche Erhebungsbezirke in die Stichprobe gelangen werden, würde dadurch letztlich die Stichprobensteuerung erschwert, und zwar dergestalt, dass ex ante bestenfalls nur ansatzweise ein vorab definierter Gesamtstichprobenumfang von Hausumringen sicherzustellen gewesen wäre. Ein vorab festgelegter Gesamtstichprobenumfang (in Höhe von 100.000 Hausumringen) war aber aus kalkulatorischen Gründen notwendig.

möglichst 50.000 oder mehr mutmaßlich GEG-relevante Nichtwohngebäude zu identifizieren sind, und der weiteren Vorgabe von identischen erhebungsbezirksspezifischen Hausumring-Stichprobenumfängen leitet sich diese Quantifizierung aus der Zahl der zu ziehenden Erhebungsbezirke ab, d. h. aus dem vorzusehenden Umfang der Erhebungsbezirksstichprobe.

Die Festlegung der Zahl zu ziehender Erhebungsbezirke steht dabei im Spannungsfeld mehrerer Ziele erhebungspraktischer und stichprobenmethodischer Art. So legte das Ziel, den bei einer zweistufigen Erhebung auftretenden nachteiligen Klumpeneffekt<sup>38</sup> weitestgehend zurückzudrängen, einen möglichst großen Erhebungsbezirksstichprobenumfang nahe. Auch in inhaltlicher Hinsicht ist ein großer Stichprobenumfang in Bezug auf Erhebungsbezirke erstrebenswert, denn mit dem Umfang der Erhebungsbezirksstichprobe verbessert sich unter sonst gleichen Bedingungen die deutschlandweite Flächenabdeckung, wodurch die ganze Vielfalt an Nichtwohngebäudetypen leichter Eingang in den im Screening identifizierten Bestand an Untersuchungsgebäuden findet. Erhebungspraktische, insbesondere erhebungsorganisatorische Erwägungen sprechen dagegen eher für eine nicht allzu große Zahl von Stichprobenerhebungsbezirken, denn dann kann der zu rekrutierende, zu schulende und zu betreuende Stamm an Screeningpersonal kleiner ausfallen. Wie im vorangegangenen Absatz ausgeführt, bedeutet ein kleiner Stichprobenumfang von Erhebungsbezirken allerdings auch, dass diese bei der vorgegebenen Zahl von 100.000 zu screenenden Hausumringen c.p. flächenmäßig größer ausfallen müssen und die Wegekosten entsprechend steigen, da sicherzustellen ist, dass auch bei einer kleineren Zahl zu ziehender Erhebungsbezirke möglichst 50.000 mutmaßlich GEG-relevante Nichtwohngebäude oder mehr gescreent werden können.

Der beschriebene Zielkonflikt wurde durch die Entscheidung zugunsten eines Stichprobenumfangs von 500 Erhebungsbezirken aufgelöst, was bei 100.000 zu screenenden Hausumringen einen erhebungsbezirksspezifischen Hausumring-Stichprobenumfang von jeweils „nur“ 200 impliziert. Bei dieser zahlenmäßigen Festlegung können die Erhebungsbezirke verhältnismäßig eng geschnitten und die Wegekosten gering gehalten werden. Gleichzeitig wird der Klumpeneffekt stark zurückgedrängt und eine befriedigende Flächenabdeckung kann erzielt werden, ohne dass der Aufwand für die Rekrutierung, Schulung und Betreuung des Screeningpersonals überhandnimmt.

Mit der Entscheidung, eine Stichprobe von 500 Erhebungsbezirken zu ziehen, folgte für die Erhebungsbezirksbildung zunächst einmal die „technische“ Notwendigkeit, wonach jeder Erhebungsbezirk mindestens 200 Hausumringe zu umfassen hat. Allerdings waren noch weitere Bedingungen zu beachten, da die Forderung nach einer „ausreichenden“ Erhebungsbezirksgröße nicht auf die Zahl eingebundener Hausumringe, sondern auf den Bestand an GEG-relevanten Nichtwohngebäuden gerichtet ist. Da ex ante, d. h. vor dem Screening, jedoch noch keine Gebäude identifiziert waren, mussten Bedingungen ersatzweise in Bezug auf Hausumringe formuliert werden. Konkret wurden folgende Bedingungen vorgegeben:

<sup>38</sup> Allgemein versteht man unter dem Klumpeneffekt den negativen Effekt einer Vergrößerung des Standardfehlers der Schätzer, der gegenüber einer uneingeschränkten Zufallsauswahl im Regelfall dann auftritt, wenn derselbe Stichprobenumfang über einen mehrstufigen Ziehungsansatz realisiert wird (vgl. (Buttler, Fickel 2002 S. 164ff.), (Stenger 1986 S. 152ff.)).



- Bedingung 1: Im Erhebungsbezirk muss es mindestens 200 Hausumringe mit Relevanzwahrscheinlichkeiten oberhalb eines bundeslandspezifisch festgesetzten Schwellenwerts geben.
- Bedingung 2: Die Summe der Relevanzwahrscheinlichkeiten muss mindestens 140 betragen, wobei nur Relevanzwahrscheinlichkeiten gezählt werden, die den bundeslandspezifisch festgesetzten Schwellenwert übersteigen.
- Bedingung 3: Im Erhebungsbezirk müssen mindestens 100 Klumpen von Hausumringen vorhanden sein, wobei in die Zählung nur solche Klumpen eingehen, denen mindestens ein Hausumring mit Relevanzwahrscheinlichkeit oberhalb des bundeslandspezifisch festgesetzten Schwellenwerts angehört.

Die erste Bedingung ist eine Verschärfung der eingangs erwähnten „technischen“ Notwendigkeit, wonach dem Erhebungsbezirk wenigstens so viele Hausumringe angehören müssen, wie gezogen werden sollen. Diese Verschärfung lässt sich mit der andernorts erörterten Übererfassung der Auswahlgrundlage begründen (vgl. Unterkapitel 3.1.3), die sich darin äußert, dass die mittlere Relevanzwahrscheinlichkeit in der Auswahlgrundlage nicht bei 1 (oder wenigstens nahe 1), sondern nur bei knapp 0,1 liegt und die Streuung dabei beachtlich ist. Die in Bedingung 1 erfolgte Verschärfung stellt sicher, dass es im Erhebungsbezirk genügend Hausumringe gibt, die zumindest mit einer Mindestwahrscheinlichkeit zu einem GEG-relevanten Nichtwohngebäude gehören. Der Schwellenwert für diese Mindestwahrscheinlichkeit wurde dabei bundeslandspezifisch festgelegt, da die Verteilung der Relevanzwahrscheinlichkeiten sowie die mittlere Relevanzwahrscheinlichkeit von Bundesland zu Bundesland schwanken, ohne dass diese Schwankungen ausschließlich mit Unterschieden im Bestand an GEG-relevanten Nichtwohngebäuden erklärbar sein dürften<sup>39</sup>.

Die zweite Bedingung stellt unter den Annahmen einer 1:1-Beziehung zwischen Hausumring und Gebäude und richtig zugewiesener bzw. unverzerrt binär-logistisch geschätzter Relevanzwahrscheinlichkeiten sicher, dass dem Erhebungsbezirk mindestens 140 GEG-relevante Nichtwohngebäude angehören und dass diese Mindestmenge allein über diejenigen Hausumringe „erreicht“ werden kann, deren Relevanzwahrscheinlichkeiten den bundeslandspezifisch festgesetzten Schwellenwert überschreiten. Vor dem Hintergrund, dass in jedem Erhebungsbezirk im Zuge des Screenings der 200er Hausumring-Stichprobe im Mittel mindestens 100 mutmaßlich GEG-relevante Nichtwohngebäude identifiziert werden müssen, um das Ziel der Übergabe von 50.000 Nichtwohngebäuden an die Breiterehebung zu erreichen, beinhaltet der Wert von 140 einen zusätzlichen Puffer.

Die dritte Bedingung löst sich von der Annahme einer 1:1-Beziehung zwischen Hausumring und Gebäude. Vor dem Hintergrund der Überlegung, dass aneinander liegende und in diesem Sinne einen Klumpen bildende Hausumringe unter Umständen zu ein und demselben Gebäude gehören, versteht sich die Klumpenvorgabe zur Sicherstellung, dass es im Erhebungsbezirk eine Mindestzahl unterschiedlicher Gebäude gibt. Die zusätzliche Vorgabe, dass im betreffenden Klumpen mindestens ein Hausumring mit einer Relevanzwahrscheinlichkeit oberhalb

<sup>39</sup> Vielmehr dürften diese Schwankungen auch ein Reflex der bundeslandspezifischen Datengrundlage für den binär-logistischen Regressionsansatz zur Schätzung der Relevanzwahrscheinlichkeiten sein.

des bundeslandspezifisch festgesetzten Schwellenwerts enthalten sein muss, gewährleistet, dass das unterstellte und sich aus den klumpenzugehörigen Hausumringen zusammensetzende Gebäude mit einer Mindestwahrscheinlichkeit ein GEG-relevantes Nichtwohngebäude ist. Bezogen auf die Erhebungsbezirksfläche und die durch sie abgedeckte Zahl an Hausumringen wirkt Bedingung 3 erhebungsbezirksvergrößernd und senkt dadurch das Risiko, dass in die zu ziehende Hausumring-Stichprobe zu viele Hausumringe gelangen, die zum selben GEG-relevanten Nichtwohngebäude gehören. Dieser erhebungsbezirksvergrößernde Effekt ist aber nicht immer gewollt. Ein Beispiel sind Innenstadtlagen, wo die dort typischerweise vorzufindenden geschlossenen Blockrandbebauungen häufig einen einzigen Hausumring-Klumpen bilden, gleichzeitig aber in aller Regel aus mehreren Gebäuden bestehen. Um unnötig große Erhebungsbezirke in innerstädtischen Lagen zu vermeiden, wurden daher flächenmäßig zu große Hausumring-Klumpen, d. h. Klumpen, deren über alle zugehörigen Hausumringe aufsummierte Grundfläche zu groß für ein „gewöhnliches“ Gebäude ist<sup>40</sup>, bei der Anwendung der in Bedingung 3 enthaltenen Zählregel fiktiv geteilt<sup>41</sup>. Unter Berücksichtigung dieser Zählregel führt Bedingung 3 somit zu einer gebremsten Erhebungsbezirksvergrößerung.

Während die erläuterten drei Bedingungen „ausreichend große“ Erhebungsbezirke bewirken sollten, traten noch zwei weitere Bedingungen für die Erhebungsbezirksbildung hinzu:

- Bedingung 4: Die Erhebungsbezirksgrenzen dürfen zusammenhängende Bebauungen nicht trennen, sondern sollen entlang von Barrieren verlaufen, die aus rechtlichen und/oder natürlichen Gründen nicht mit Gebäuden bebaubar sind (z. B. Straßen, Flüsse).
- Bedingung 5: Bundeslandgrenzen sind zu wahren, d. h. ein Erhebungsbezirk darf sich nicht auf mehr als ein Bundesland erstrecken.

Bedingung 4 stellt sicher, dass ein und dasselbe Gebäude nicht in mehr als einem Erhebungsbezirk liegen kann. Sie beugt dadurch erfassungstechnische Probleme vor Ort vor<sup>42</sup> und erleichtert die an späterer Stelle erörterte Berechnung gebäudebezogener Ziehungswahrscheinlichkeiten (vgl. Unterkapitel 3.2.4).

Wie bereits ausgeführt, ist Bedingung 5 dem Umstand geschuldet, dass die Verteilungsmuster der Relevanzwahrscheinlichkeiten eine Bundeslandabhängigkeit aufweisen.

Im Zuge der durch den Projektpartner IÖR vorgenommenen Erhebungsbezirksbildung wurde das deutsche Staatsgebiet flächendeckend und überlappungsfrei in 7.465 Erhebungsbezirke aufgeteilt<sup>43</sup>. Die Spannweite der Erhebungsbezirksflächen

<sup>40</sup> Als Grenzwert wurde 500 qm angesetzt.

<sup>41</sup> War die im ungeteilten Klumpen vorhandene Zahl der Hausumringe mit Relevanzwahrscheinlichkeit oberhalb des Schwellenwerts kleiner als die Zahl der fiktiven „Teilkumpen“, wurde zur Prüfung von Bedingung 3 der ungeteilte Klumpen mit eben dieser Zahl an Hausumringen gezählt, andernfalls wurde der ungeteilte Klumpen entsprechend seiner (nicht notwendigerweise ganzzahligen) Zahl fiktiver „Teilkumpen“ gewertet.

<sup>42</sup> Andernfalls hätte sich für die Screener die Notwendigkeit ergeben können, bei der Umsetzung der Gebäudedefinition Hausumringe aus anderen Erhebungsbezirken ergänzen zu müssen. Die zur Gebäudeerfassung und Datenaufnahme benutzte mobile App war jedoch so konfiguriert, dass sie nur Hausumringe desjenigen Erhebungsbezirks einblendete, für den der Screener eingesetzt wurde.

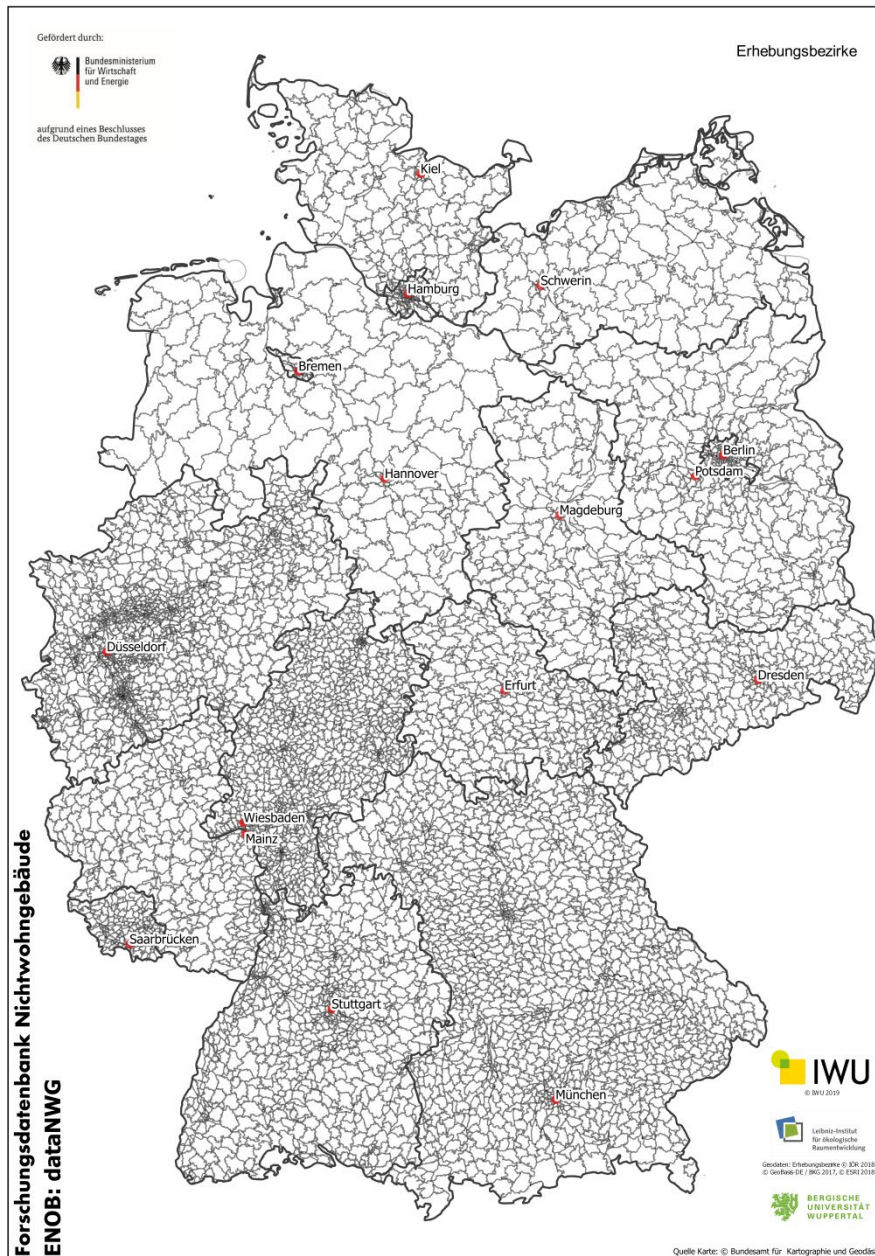
<sup>43</sup> Der hierzu angewandte Algorithmus wird in (Schorcht et al. 2021) ausführlich beschrieben.



reicht dabei von 0,11 km<sup>2</sup> bis 1.003 km<sup>2</sup> bei einer mittleren Fläche in Höhe von 48,07 km<sup>2</sup>. Wie Abbildung 3-13 zu entnehmen ist, zeigt sich in der Tendenz das erwartete Bild von flächenmäßig kleineren Bezirken in wirtschaftsstarken Agglomerationsräumen sowie größeren Erhebungsbezirksflächen im ländlichen Raum. Darüber hinaus hängen die Flächen gebildeter Erhebungsbezirke von bundeslandspezifischen Besonderheiten im Katasterwesen ab, konkret von Art und Aussagefähigkeit der Nutzungsattribute aus dem Amtlichen Liegenschaftskatasterinformationssystem (ALKIS). Denn wie in Unterkapitel 3.1.3 dargestellt, wurden die vorgefundenen Nutzungsattribute herangezogen, um Hausumringe hinsichtlich der Eigenschaft, zu einem GEG-relevanten Nichtwohngebäude zu gehören, in Gestalt von (Relevanz-) Wahrscheinlichkeiten bewertet, die wiederum für die Erhebungsbezirksbildung maßgeblich waren. So lassen sich die auffällig großen Erhebungsbezirksflächen in Niedersachsen – die mittlere Erhebungsbezirksfläche liegt in diesem Bundesland bei beachtlichen 247,87 km<sup>2</sup> – auch darauf zurückführen, dass in diesem Bundesland nur 0,4% aller Hausumringe als eindeutig zu einem GEG-relevanten Nichtwohngebäude gehörend (zugewiesene Relevanzwahrscheinlichkeit: 1,0) identifiziert wurden und die mittlere Relevanzwahrscheinlichkeit nur knapp 0,039 beträgt, während im Bundesdurchschnitt die entsprechenden Werte bei 1% bzw. 0,097 liegen.

Was die Zahl der erhebungsbezirkszugehörigen Hausumringe angeht, liegt der Mittelwert bei 6.536. Der in dieser Hinsicht kleinste Erhebungsbezirk umfasst 233 Hausumringe und liegt in Hamburg, dem am anderen Ende der Skala ein Erhebungsbezirk mit 70.348 Hausumringen in Niedersachsen gegenübersteht.

**Abbildung 3-13: Die Aufteilung Deutschlands in Erhebungsbezirke**



Quelle: Eigene Darstellung

### Schichtung und sich daran anschließende Ziehung

Wie in Kapitel 2 ausgeführt, sollen aus allen Landesteilen ausreichend viele Nichtwohngebäude in der Stichprobe vertreten sein. Das Mittel der Wahl zur Erfüllung dieser Anforderung besteht in der Ziehung einer geschichteten Stichprobe (*stratified sampling*) von Erhebungsbezirken unter Zugrundelegung eines geeigneten räumlichen Schichtungskriteriums, also darin, die Gesamtheit aller 7.465 Erhebungsbezirke räumlich zu gruppieren und aus jeder Gruppe bzw. Schicht (*stratum*)

eine eigene Erhebungsbezirksstichprobe zu ziehen. Als räumliche Schichtungskriterien wurden dazu die Bundeslandzugehörigkeit und unterhalb der Bundeslandebene – soweit sinnvoll<sup>44</sup> – die Raumordnungsregion herangezogen<sup>45</sup>. Insgesamt wurden auf diese Weise 96 Schichten gebildet.

**Tabelle 3-4: Bundeslandspezifische Erhebungsbezirksstichprobenumfänge**

| Bundesland                   | Anzahl Hausumringe | Anteil Hausumringe | Gesamtzahl Erhebungsbezirke | Stichprobenumfang Erhebungsbezirke |
|------------------------------|--------------------|--------------------|-----------------------------|------------------------------------|
| Schleswig-Holstein           | 1.933.556          | 4,0%               | 198                         | 20                                 |
| Freie und Hansestadt Hamburg | 361.453            | 0,7%               | 121                         | 4                                  |
| Niedersachsen                | 5.420.848          | 11,1%              | 195                         | 56                                 |
| Freie Hansestadt Bremen      | 241.010            | 0,5%               | 36                          | 2                                  |
| Nordrhein-Westfalen          | 9.385.191          | 19,2%              | 1.484                       | 96                                 |
| Hessen                       | 4.049.719          | 8,3%               | 1.134                       | 42                                 |
| Rheinland-Pfalz              | 2.867.928          | 5,9%               | 309                         | 29                                 |
| Baden-Württemberg            | 5.661.895          | 11,6%              | 882                         | 58                                 |
| Freistaat Bayern             | 8.330.471          | 17,1%              | 1.504                       | 84*                                |
| Saarland                     | 565.062            | 1,2%               | 152                         | 6                                  |
| Berlin                       | 508.801            | 1,0%               | 138                         | 5                                  |
| Brandenburg                  | 2.333.686          | 4,8%               | 287                         | 24                                 |
| Mecklenburg-Vorpommern       | 1.129.249          | 2,3%               | 149                         | 12                                 |
| Freistaat Sachsen            | 2.329.689          | 4,8%               | 368                         | 24                                 |
| Sachsen-Anhalt               | 1.666.516          | 3,4%               | 189                         | 17                                 |
| Freistaat Thüringen          | 2.006.226          | 4,1%               | 319                         | 21                                 |
| <b>gesamt</b>                | <b>48.791.300</b>  | <b>100%</b>        | <b>7.465</b>                | <b>500</b>                         |

\* Rechnerisch ergibt sich für Bayern ein – gerundeter – Stichprobenumfang von 85 Erhebungsbezirken. Um die Vorgabe eines Gesamtstichprobenumfangs von 500 Erhebungsbezirken exakt zu erfüllen, wurde der für Bayern festgelegte Stichprobenumfang um einen Erhebungsbezirk auf 84 abgesenkt.

Quelle: Eigene Berechnung

<sup>44</sup> In den Stadtstaaten Berlin und Hamburg sowie im Saarland fällt die Raumordnungsregion mit den Landesgrenzen zusammen, so dass sich eine weitere Gruppierung der Erhebungsbezirke unterhalb der Landesgrenze erübrigt. Im Bundesland Bremen bildet zwar jeder der beiden landeszugehörigen Großstädte (Bremen und Bremerhaven) eine eigene Raumordnungsregion. Angesichts des sehr geringen Stichprobenumfangs von nur zwei Erhebungsbezirken aus dem Bundesland Bremen (vgl. Tabelle 3-4) ist eine Schichtung nach Raumordnungsregionen jedoch nicht sinnvoll.

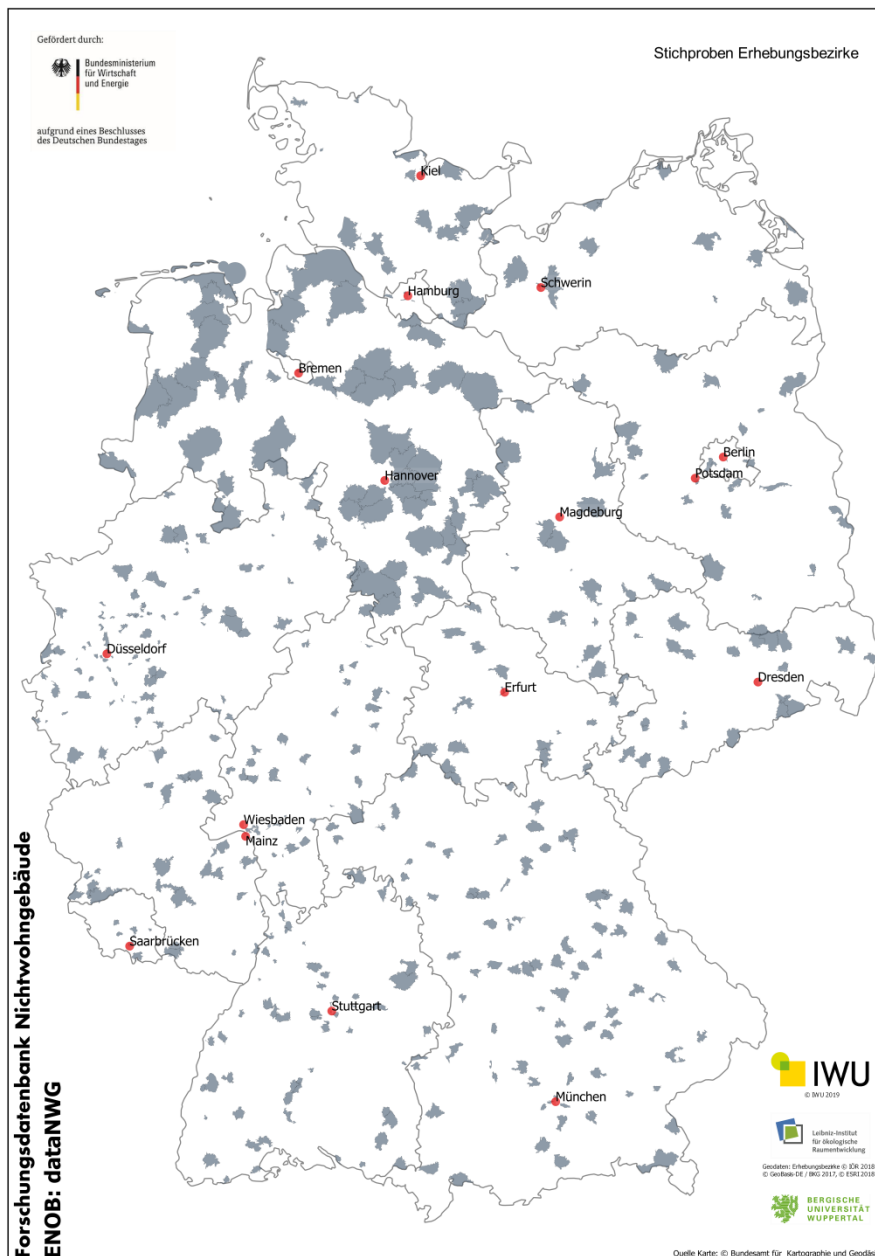
<sup>45</sup> Da bei der Bildung der Erhebungsbezirke nicht auf deren raumordnungsscharfe Abgrenzung geachtet wurde, erstrecken sich einige Erhebungsbezirke auf mehr als eine Raumordnungsregion. Für die zur Schichtenbildung erforderliche Zuordnung zu einer einzigen Raumordnungsregion wurde in diesen Fällen auf den innenliegenden Mittelpunkt (entspricht i.d.R. dem Schwerpunkt) des Erhebungsbezirks zurückgegriffen (vgl. (Schorcht et al. 2021)).

Die Festlegung der schichtenspezifischen Stichprobenumfänge unter der Maßgabe von insgesamt 500 zu ziehenden Erhebungsbezirken erfolgte in zwei Schritten. Im ersten Schritt wurde der bundeslandspezifische Stichprobenumfang festgelegt, wobei die Aufteilung proportional zur Zahl der Hausumringe des betreffenden Bundeslandes erfolgte. Tabelle 3-4 weist in der rechten Tabellenspalte die so ermittelten bundeslandspezifischen Stichprobenumfänge aus.

Im zweiten Schritt wurde der bundeslandspezifische Stichprobenumfang auf die bundeslandzugehörigen Raumordnungsregionen heruntergebrochen, und zwar dergestalt, dass der einer Raumordnungsregion zugewiesene Stichprobenumfang proportional zur Zahl ihrer Erhebungsbezirke stand.

Die sich an die Schichtung anschließende Ziehung der Erhebungsbezirksstichprobe erfolgte uneingeschränkt zufällig ohne Zurücklegen (*simple random sampling without replacement*). Abbildung 3-14 zeigt, wie sich die gezogenen Erhebungsbezirke (in der Abbildung als graue Flächen dargestellt) im Raum verteilen, und belegt die intendierte regionale Ausgewogenheit der Erhebungsbezirksgesamtstichprobe.

**Abbildung 3-14: Räumliche Verteilung der gezogenen Erhebungsbezirke**



Quelle: Eigene Darstellung

### 3.2.3 2. Ziehungsstufe: Geschichtete Zufallsauswahl von Hausumringen

Die Stichprobenziehung von Hausumringen auf der zweiten Ziehungsstufe (d. h. innerhalb zuvor gezogener Erhebungsbezirke) soll dafür Sorge tragen, dass im Screening möglichst 50.000 oder mehr GEG-relevante Nichtwohngebäude identifiziert werden (vgl. Unterkapitel 3.2.1). Da mutmaßlich nur ein kleiner Teil der

ENOB:dataNWG

Hausumringe zu GEG-relevanten Nichtwohngebäuden gehört, ist die Stichprobenziehung so zu gestalten, dass Hausumringe, die eher zu einem GEG-relevanten Nichtwohngebäude gehören, überproportional in die Stichprobe aufgenommen werden. Das stichprobenmethodische Mittel der Wahl hierfür ist eine nach Relevanzwahrscheinlichkeiten disproportional geschichtete Stichprobenziehung, die in ihren Grundzügen nachfolgend beschrieben wird.

### **Schichtung und sich daran anschließende Ziehung**

Zur Durchführung einer geschichteten Stichprobenziehung müssen Entscheidungen hinsichtlich der Definition und der Zahl der Schichten sowie der schichtenspezifischen Stichprobenumfänge getroffen werden.

Da im vorliegenden Fall das Ziel darin besteht, Hausumringe mit höheren Relevanzwahrscheinlichkeiten überproportional in die Hausumring-Stichprobe aufzunehmen, drängt sich eine Schichtung entlang von Relevanzwahrscheinlichkeitsklassen auf. Festzulegen bleiben noch Zahl und Breite der Relevanzwahrscheinlichkeitsklassen. Was die Zahl angeht, fiel die Entscheidung auf fünf Relevanzwahrscheinlichkeitsklassen, d. h. fünf Schichten je Erhebungsbezirk. Bezüglich der Klassenbreite standen zwei Alternativen zur Wahl, nämlich fest vorgegebene Klassenbreiten für alle Erhebungsbezirke oder variable, also erhebungsbezirksindividuelle Klassenbreiten. Die Wahl fiel angesichts des Umstands, dass die Verteilung der Relevanzwahrscheinlichkeiten von Erhebungsbezirk zu Erhebungsbezirk schwankt, zugunsten variabler Klassenbreiten. Denn bei festen Klassenbreiten hätte das Risiko bestanden, dass in Erhebungsbezirken mit „extremen“ Verteilungsmustern manche Schichten unbesetzt bleiben oder dass fast alle Hausumringe in dieselbe Schicht fallen und dadurch das Schichtungsziel unterlaufen wird. Um von Erhebungsbezirk zu Erhebungsbezirk die Klassenbreiten festzulegen, kam ein sog. k-Means-Algorithmus zum Einsatz. Dieser teilt eine Menge von Objekten (hier: Hausumringe) anhand eines Ähnlichkeitskriteriums (hier: ähnliche Relevanzwahrscheinlichkeiten) in k Gruppen (hier: k = 5 Schichten) auf, wobei er Gruppen mit geringer Varianz und ähnlicher Größe bevorzugt.

Die Aufteilung des erhebungsbezirksspezifischen Gesamtstichprobenumfangs von 200 Hausumringen auf die fünf wie beschrieben erzeugten Schichten erfolgte zwar mit dem Ziel, möglichst viele Hausumringe zu ziehen, die die Identifizierung von GEG-relevanten Nichtwohngebäuden ermöglichen. Hierbei galt es aber, diverse Randbedingungen zu beachten.

So war aus stichprobenmethodischen Gründen zu gewährleisten, dass in jeder der fünf Schichten mindestens zwei Hausumringe gezogen werden<sup>46</sup>.

Darüber hinaus sollten die Ziehungswahrscheinlichkeiten der Hausumringe nicht unnötig stark schwanken. Auf keinen Fall sollten die Ziehungswahrscheinlichkeiten<sup>47</sup> unter 1/300 fallen. Die Vorgabe einer solchen Untergrenze erfolgte mit Blick auf die statistische Belastbarkeit späterer Ergebnisauswertungen und vor dem Hintergrund einer angesichts von Abbildung 3-10 zu erwartenden rechtsschiefen

<sup>46</sup> Andernfalls hätte nicht sichergestellt werden können, dass zwei beliebige Nichtwohngebäude in die Stichprobe gelangen können. Diese Voraussetzung ist aber notwendig, um approximativ erwartungstreue Schätzungen des Standardfehlers zu ermöglichen (vgl. Kapitel 2).

<sup>47</sup> Gemeint ist die bedingte Hausumring-Ziehungswahrscheinlichkeit, d. h. die Ziehungswahrscheinlichkeit unter der Bedingung, dass auf der vorgelagerten ersten Ziehungsstufe der betreffende Erhebungsbezirk gezogen wird.



Verteilung der Relevanzwahrscheinlichkeiten in den einzelnen Erhebungsbezirken. Denn diese Verteilungseigenschaft bewirkt, dass die jeweils unterste Schicht, also diejenige Schicht, in der die Hausumringe mit den geringsten Relevanzwahrscheinlichkeiten zu finden sind, im Regelfall besonders stark besetzt ist. Würden nun nur wenige (im Extremfall nur zwei) Hausumringe aus solchen Schichten gezogen werden und würde sich im Nachhinein herausstellen, dass diese Hausumringe wider Erwarten doch zu GEG-relevanten Nichtwohngebäuden gehören, bekämen diese Gebäude weit überdurchschnittlich große Hochrechnungsfaktoren zugewiesen<sup>48</sup> und könnten daher die späteren Auswertungsergebnisse stark beeinflussen. Dies gilt umso mehr für teilsegmentspezifische Auswertungen, die angesichts der Heterogenität des Nichtwohngebäudebestands von besonderem Interesse sind. Die Vorgabe einer Mindestziehungswahrscheinlichkeit von 1/300 wirkt somit allzu großen Hochrechnungsfaktoren mitsamt den beschriebenen negativen Auswirkungen entgegen.

Ausgangspunkt für die Ermittlung schichtenspezifischer Auswahlätze war in jedem Erhebungsbezirk eines Bundeslandes<sup>49</sup> zunächst eine proportionale Aufteilung des erhebungsbezirksspezifischen Gesamtstichprobenumfangs (in Höhe von 200 Hausumringen), wodurch jeder Hausumring des Erhebungsbezirks ungeachtet seiner Schichtenzugehörigkeit dieselbe Ziehungswahrscheinlichkeit aufweisen würde. Danach wurden in allen Erhebungsbezirken des Bundeslandes die schichtenspezifischen Stichprobenumfänge unter Maßgabe der obigen Randbedingungen „von unten nach oben“ umverteilt, um das Schichtungsziel, nämlich die überproportionale Ziehung von Hausumringen mit hohen Relevanzwahrscheinlichkeiten, zu erreichen. Konkret wurden die den jeweils unteren beiden Schichten zugewiesenen Stichprobenumfänge<sup>50</sup> sukzessive abgesenkt und der obersten Schicht, d. h. derjenigen Schicht mit den höchsten Relevanzwahrscheinlichkeiten, zugeordnet. Entstand im Laufe dieser Umverteilungsprozedur dabei die Situation, dass in der obersten Schicht der Stichprobenumfang mit der Besetzungstärke der Schicht gleichzog, verlagerte sich die weitere Umverteilungsprozedur auf die zweitoberste Schicht und – falls auch in dieser der Stichprobenumfang so hoch war wie die Anzahl der schichtenzugehörigen Hausumringe – auf die drittoberste, also mittlere Schicht. Dieser Umverteilungsalgorithmus stoppte spätestens<sup>51</sup> dann, wenn vorgegebene Schwellenwerte erreicht wurden, und zwar Schwellenwerte in Bezug auf die bei einer Stichprobenziehung von insgesamt 200 Hausumringen zu erwartende Summe der Relevanzwahrscheinlichkeiten. Die zu erwartende Summe der

<sup>48</sup> Der (Gebäude-) Hochrechnungsfaktor entspricht dem Kehrwert der (Gebäude-) Ziehungswahrscheinlichkeit und gibt an, für wie viele Gebäude der Grundgesamtheit das Stichprobengebäude stellvertretend steht.

<sup>49</sup> Das nachfolgend beschriebene Procedere erfolgte mit Blick auf Unterschiede im Katasterwesen, die sich nicht zuletzt in unterschiedlichen Verteilungsstrukturen der Relevanzwahrscheinlichkeiten und in Unterschieden im Verteilungsmuster der Hausumring-Gebäude-Beziehungsgeflechte manifestieren, nach Bundesländern getrennt.

<sup>50</sup> Die Entscheidung, in welchen der beiden unteren Schichten der Stichprobenumfang reduziert wird, wurde davon abhängig gemacht, wie sich die Ziehungswahrscheinlichkeiten nach Reduzierung des Stichprobenumfangs darstellen. Konkret fiel die Entscheidung zugunsten der Schicht mit der höheren „neuen“ Ziehungswahrscheinlichkeit. Hintergrund für diese Entscheidung war das Bestreben, die Streuung der Ziehungswahrscheinlichkeiten zu begrenzen, um der beschriebenen Problematik zu großer gebäudespezifischer Hochrechnungsfaktoren gegenzusteuern. Waren die neuen Ziehungswahrscheinlichkeiten ausnahmsweise identisch, wurde der Stichprobenumfang der untersten Schicht reduziert.

<sup>51</sup> Der Umverteilungsalgorithmus stoppte vorher, wenn in einer der beiden unteren Schichten der Stichprobenumfang auf eins hätte abgesenkt werden müssen oder wenn die Absenkung des Stichprobenumfangs Ziehungswahrscheinlichkeiten unterhalb des Grenzwerts von 1/300 zur Folge gehabt hätte.

Relevanzwahrscheinlichkeiten berechnete sich dabei, indem in jeder der fünf Schichten das Produkt aus mittlerer Relevanzwahrscheinlichkeit und vorgesehendem Stichprobenumfang gebildet wurde und anschließend die so gebildeten Produkte aufsummiert wurden. Was die Schwellenwerte angeht, wurden neun Grenzwerte geprüft: 100, 110, 120, 130, 140, 150, 160, 170 und 180. Unter den Annahmen verzerrungsfrei geschätzter Relevanzwahrscheinlichkeiten sowie einer 1:1-Beziehung zwischen Hausumring und Gebäude geben diese Schwellenwerte somit die zu erwartende Zahl von GEG-relevanten Nichtwohngebäuden an, die durch die Stichprobenziehung von Hausumringen erreicht werden. Um eine Entscheidung hinsichtlich des bundeslandspezifisch anzulegenden Schwellenwerts zu treffen, kam ein sog. Monte-Carlo-Experiment zum Einsatz. Konkret wurde für jeden der neun infrage kommenden Schwellenwerte das zweistufige Ziehungsprozedere, bestehend aus Erhebungsbezirksstichprobenziehung mit anschließender Ziehung einer geschichteten Hausumring-Stichprobe, im betreffenden Bundesland 50 Mal wiederholt und für jede gezogene Stichprobe wurde analysiert, mit wie vielen GEG-relevanten Nichtwohngebäuden im ungünstigsten Fall<sup>52</sup> zu rechnen ist. Diese Anzahl wurde dann über die 50 Simulationsläufe gemittelt. Die Entscheidung fiel dann grundsätzlich zugunsten desjenigen Schwellenwerts, der in diesem Worst-Case-Szenario dahingehend am besten abschnitt, dass er im Mittel über die 50 Simulationsläufe die größte Zahl GEG-relevanter Nichtwohngebäude versprach, ohne dass dieser Mittelwert von Simulationslauf zu Simulationslauf zu stark schwankte.

Zu betonen ist, dass die Entscheidung zugunsten des Schwellenwerts aus stichprobenmethodischen Gründen<sup>53</sup> vor der Stichprobenziehung der Erhebungsbezirke erfolgte.

Die sich an die Schichtung anschließende Ziehung der Hausumring-Stichprobe unter Zugrundelegung der wie beschrieben ermittelten schichtenspezifischen Stichprobenumfänge erfolgte – wie auch die Ziehung der Erhebungsbezirke auf der vorgelagerten ersten Ziehungsstufe – uneingeschränkt zufällig ohne Zurücklegen (*simple random sampling without replacement*).

Das dargestellte Schichtungs- und Ziehungsprozedere soll am Beispiel eines Berliner Erhebungsbezirks dargestellt werden. Aus darstellerischen Gründen handelt es sich dabei um einen nur knapp 1,0 km<sup>2</sup> großen Erhebungsbezirk in Berlin Mitte. Wie Tabelle 3-5 zu entnehmen ist, umfasst der Erhebungsbezirk 390 Hausumringe<sup>54</sup>, die durch den k-Means-Algorithmus auf k = 5 Hausumring-Schichten unterschiedlicher Breite aufgeteilt wurden. Erwartungsgemäß ist die unterste Schicht (Schicht 1), d. h. diejenige Schicht mit den niedrigsten

<sup>52</sup> Vereinfacht gesprochen ist der ungünstigste Fall dadurch gekennzeichnet, dass die 1:1-Beziehung zwischen Hausumring und Gebäude bei zusammenklumpenden Hausumringen nicht gilt, sondern dass stattdessen alle Hausumringe eines Klumpens zu ein und demselben Gebäude gehören. In einem solchen Szenario ist die Gefahr gegeben, dass mehrere Stichproben-Hausumringe zur Identifizierung desselben Gebäudes führen, was dem Ziel, im Screening möglich viele (unterschiedliche) GEG-relevante Nichtwohngebäude zu identifizieren, abträglich ist.

<sup>53</sup> Anderweitig wäre die Annahme der Invarianz nicht erfüllt, also die Annahme, dass das Stichprobendesign auf einer nachgelagerten Ziehungsstufe unabhängig davon ist, welche Erhebungseinheiten der vorgelagerten Ziehungsstufe ausgewählt werden (Särndal et al. 1992 S. 134). Die Annahme der Invarianz vereinfacht die formelmäßige Beschreibung des Stichprobendesigns, die in den nachfolgenden Kapiteln und Unterkapiteln entwickelt wird.

<sup>54</sup> Gemessen an der Zahl der Hausumringe ist dieser Erhebungsbezirk der viertkleinste, gemessen an der Fläche sogar der kleinste der 138 Berliner Erhebungsbezirke.



Relevanzwahrscheinlichkeiten, mit 159 Hausumringen am stärksten besetzt. Im Falle einer Aufteilung des erhebungsbezirksspezifischen Stichprobenumfangs (200 Hausumringe) proportional zur Besetzungsstärke der Schichten ergäbe sich für Schicht 1 ein Stichprobenumfang von 81 und für Schicht 5 von 64. Im Falle der Ziehung der Hausumring-Stichprobe unter Anlegung dieser schichtenspezifischen Stichprobenumfänge liegt der Erwartungswert der über die Stichproben-Hausumringe aufsummierten Relevanzwahrscheinlichkeiten bei gerade einmal ca. 103. Die Berechnung von Worst-Case-Szenarien legte für Berlin jedoch nahe, einen Wert von 120 anzustreben. Das dazu erforderliche Umverteilungsverfahren führte im betrachteten Erhebungsbezirk zu der in der drittuntersten Tabellenzeile dargestellten Endverteilung der schichtenspezifischen Stichprobenumfänge. Verglichen mit der Anfangsverteilung wurden dazu die Stichprobenumfänge von Schicht 1 und 2 um insgesamt 17 Fälle reduziert und der Stichprobenumfang für Schicht 5 wurde um dieselbe Zahl von Fällen angehoben. Folgerichtig sind die Ziehungswahrscheinlichkeiten nicht mehr identisch, sondern haben sich zugunsten der Hausumringe in Schicht 5 verschoben. In der untersten Tabellenzeile sind die resultierenden (hausumringspezifischen) Hochrechnungsfaktoren ausgewiesen. Sie reichen dabei von 1,53 in Schicht 5 bis zu 2,45 in Schicht 1.

**Tabelle 3-5: Festlegung der schichtenspezifischen Stichprobenumfänge in einem Beispiel-Erhebungsbezirk („BE 70“) in Berlin Mitte**

|                                                            | Schicht 1 | Schicht 2 | Schicht 3 | Schicht 4 | Schicht 5 | gesamt |
|------------------------------------------------------------|-----------|-----------|-----------|-----------|-----------|--------|
| mittlere Relevanzwahrscheinlichkeit                        | 0,01      | 0,29      | 0,61      | 0,95      | 1,00      | 0,51   |
| Besetzungsstärke                                           | 159       | 20        | 43        | 44        | 124       | 390    |
| Anfangsverteilung: Stichprobenumfang bei prop. Aufteilung  | 81        | 10        | 22        | 23        | 64        | 200    |
| Endverteilung des Stichprobenumfangs                       | 65        | 9         | 22        | 23        | 81        | 200    |
| hausumringspez. Ziehungswahrscheinlichkeit (= Auswahlatz)* | 0,41      | 0,45      | 0,51      | 0,52      | 0,65      | -      |
| hausumringspez. Hochrechnungsfaktor*                       | 2,45      | 2,22      | 1,95      | 1,91      | 1,53      | -      |

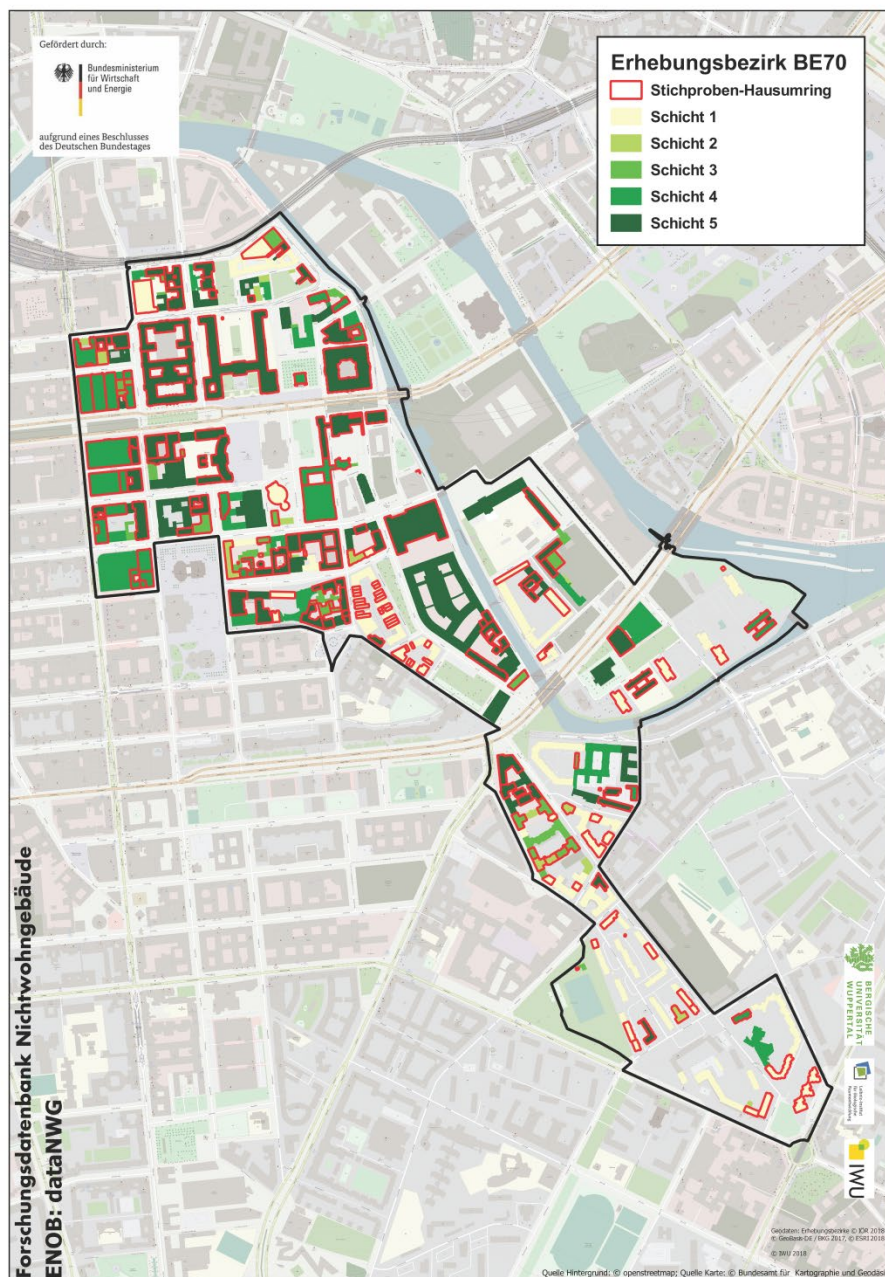
\* jeweils ohne Berücksichtigung der Ziehungswahrscheinlichkeit für den Erhebungsbezirk

Quelle: Eigene Berechnung

Beim in Tabelle 3-5 dargestellten Erhebungsbezirk handelt es sich um einen der fünf tatsächlich gezogenen Berliner Erhebungsbezirke. Das Procedere zur Festlegung der schichtenspezifischen Stichprobenumfänge erfolgte in diesem Erhebungsbezirk daher nicht nur zur Berechnung von Worst-Case-Szenarien zwecks Festlegung des berlinspezifischen Schwellenwerts (von 120). Vielmehr waren die bei diesem Schwellenwert abgeleiteten und in Tabelle 3-5 aufgeführten schichtenspezifischen Stichprobenumfänge auch maßgebend für die eigentliche (finale) Stichprobenziehung. Die so gezogenen 200 Hausumringe sind in Abbildung 3-15, ENOB:dataNWG

die den gesamten Erhebungsbezirk mitsamt all seiner Hausumringe kartografisch darstellt, durch rote Umrandung besonders gekennzeichnet. Wie aus der Abbildung ersichtlich, verteilen sich die 200 gezogenen Hausumringe auf alle fünf farblich voneinander abgesetzten Hausumring-Schichten, die entlang von Relevanzwahrscheinlichkeiten mittels des k-Means-Algorithmus gebildet wurden.

**Abbildung 3-15:** Kartografische Darstellung des Beispiel-Erhebungsbezirks „BE 70“ von Tabelle 3-5



Quelle: Eigene Darstellung

### 3.2.4 Ziehungswahrscheinlichkeiten

Das Stichprobenauswahlverfahren manifestiert sich letztlich in gebäudespezifischen Ziehungswahrscheinlichkeiten, auf die Schätzungen von populationsbeschreibenden Parametern ebenso zurückgreifen wie Schätzungen der korrespondierenden Standardfehler. Zur Ableitung gebäudespezifischer Ziehungswahrscheinlichkeiten ist im vorliegenden Fall dabei nicht nur das zweistufige Auswahlverfahren zu berücksichtigen, sondern auch der Umstand des potentiellen Auseinanderfallens von Hausumringen (Erhebungseinheiten) und Nichtwohngebäuden (Untersuchungseinheiten). Darüber hinaus sind zwei Arten von Ziehungswahrscheinlichkeiten zu unterscheiden, nämlich die sog. Inklusionswahrscheinlichkeiten 1. und die 2. Ordnung.

Die Inklusionswahrscheinlichkeit 1. Ordnung stellt – ganz allgemein ausgedrückt – auf die Wahrscheinlichkeit ab, dass ein Ziehungsobjekt in die Stichprobe gelangt. Inklusionswahrscheinlichkeiten 1. Ordnung sind somit das, was gemeinhin unter Ziehungswahrscheinlichkeiten verstanden wird und in diesem Sinne wurde und wird in diesem Bericht von Ziehungswahrscheinlichkeiten anstelle von Inklusionswahrscheinlichkeiten 1. Ordnung gesprochen. Im Rahmen des vorliegenden Projekts gibt es drei Arten von Ziehungsobjekten und damit auch drei Arten von Inklusionswahrscheinlichkeiten 1. Ordnung bzw. drei Arten von Ziehungswahrscheinlichkeiten. Ziehungsobjekte sind konkret Erhebungsbezirke, Hausumringe und Nichtwohngebäude. Dass Nichtwohngebäude dabei anders als Erhebungsbezirke und Hausumringe nicht direkt gezogen werden, sondern nur indirekt (über gezogene Hausumringe) in die Stichprobe gelangen, ändert nichts daran, dass auch Nichtwohngebäude Ziehungswahrscheinlichkeiten aufweisen.

Demgegenüber sind Inklusionswahrscheinlichkeiten 2. Ordnung kombinierte Ziehungswahrscheinlichkeiten, nämlich Wahrscheinlichkeiten, dass zwei verschiedene Ziehungsobjekte gemeinsam in die Stichprobe aufgenommen werden.

Während zur Schätzung populationsbeschreibender Parameter ausschließlich Inklusionswahrscheinlichkeiten 1. Ordnung erforderlich sind, bedingen Standardfehler-schätzungen beide Arten von Inklusionswahrscheinlichkeiten.

Im Folgenden werden zunächst die „konventionellen“ Ziehungswahrscheinlichkeiten, also die Inklusionswahrscheinlichkeiten 1. Ordnung, vorgestellt, bevor die deutlich komplexer aufgebauten Inklusionswahrscheinlichkeiten 2. Ordnung erörtert werden.

#### **Inklusionswahrscheinlichkeiten 1. Ordnung („Ziehungswahrscheinlichkeiten“)**

Auch wenn es im vorliegenden Projekt – wie gerade ausgeführt – drei Arten von Ziehungsobjekten gibt, interessieren die Ziehungswahrscheinlichkeiten von Nichtwohngebäuden in besonderer Weise, da Nichtwohngebäude als Untersuchungseinheiten fungieren. Aufgrund dessen konzentriert sich die nachfolgende Darstellung auf die Ermittlung gebäudebezogener Ziehungswahrscheinlichkeiten.

Die Ziehungswahrscheinlichkeit für eine zur primären Erhebungseinheit  $i$  gehörende Untersuchungseinheit  $k$ , d. h. die Ziehungswahrscheinlichkeit für ein zum Erhebungsbezirk  $i$  gehörendes Nichtwohngebäude  $k$ , wird nachfolgend mit  $\pi_k$  bezeichnet und entspricht in einem zweistufigen Ziehungsverfahren dem Produkt zweier Ziehungswahrscheinlichkeiten, nämlich

- der Ziehungswahrscheinlichkeit für den Erhebungsbezirk  $i$  ( $\pi_{Ii}$ )<sup>55</sup> und
- der bedingten Ziehungswahrscheinlichkeit für das Nichtwohngebäude  $k$  unter der Bedingung der vorgelagerten Ziehung des zugehörigen Erhebungsbezirks ( $\pi_{k|i}$ ).

Es gilt daher

$$(1) \quad \pi_k = \pi_{Ii} \pi_{k|i}$$

Für die erhebungsbezirksspezifische Ziehungswahrscheinlichkeit  $\pi_{Ii}$  gilt aufgrund des Ziehens ohne Zurücklegen

$$(2) \quad \pi_{Ii} = \frac{n_I(g)}{N_I(g)} \text{ für } i \in U_{Ig} \text{ mit}$$

$U_{Ig}$  : Menge der primären Erhebungseinheiten (d. h. der Erhebungsbezirke) in Schicht  $g$

$n_I(g)$  : Stichprobenumfang (d. h. Anzahl gezogener Erhebungsbezirke) in Schicht  $g$

$N_I(g)$  : Anzahl der primären Erhebungseinheiten (d. h. der Erhebungsbezirke) in Schicht  $g$

Alle Erhebungsbezirke der Schicht  $g$  haben somit dieselbe Ziehungswahrscheinlichkeit, wobei es insgesamt 96 solcher Erhebungsbezirksschichten gibt, die mit den Raumordnungsregionen (ROR) identisch sind und auf die die 7.465 Erhebungsbezirke aufgeteilt wurden. Insgesamt wurden 500 Erhebungsbezirke gezogen (vgl. Unterkapitel 3.2.2).

Auch wenn auf der zweiten Ziehungsstufe direkt nur Hausumringe gezogen wurden und nur diese damit als sekundäre Erhebungseinheiten fungieren, interessieren mit Blick auf die späteren Ergebnisauswertungen entlang von Nichtwohngebäuden gebäudespezifische Ziehungswahrscheinlichkeiten. Dabei ist zu berücksichtigen, dass zwischen Hausumringen und (Nichtwohn-) Gebäuden nicht notwendigerweise eindeutige Beziehungen gegeben sind. Vielmehr sind entsprechend der Ausführungen in Unterkapitel 3.1.2 vier Fallkonstellationen zu unterscheiden.

In **Fallkonstellation 1** besteht eine eindeutige Beziehung zwischen dem Hausumring als sekundärer Erhebungseinheit und dem Nichtwohngebäude als Untersuchungseinheit. Aufgrund dessen entspricht die Ziehungswahrscheinlichkeit des Hausumrings der des betreffenden Nichtwohngebäudes. Für die (bedingte) Gebäudeziehungswahrscheinlichkeit  $\pi_{k|i}$  des Nichtwohngebäudes  $k$  gilt aufgrund des Ziehens ohne Zurücklegen daher

$$(3) \quad \pi_{k|i} = \frac{n_i(h)}{N_i(h)} \text{ für } k \in U_i \text{ mit}$$

$U_i$  : Menge der Untersuchungseinheiten (d. h. der Nichtwohngebäude) in Erhebungsbezirk  $i$

<sup>55</sup> Die im Index stehende und kursiv gedruckte römische Ziffer  $I$  kennzeichnet hier und im Folgenden die erste Ziehungsstufe.

- $n_i(h)$  : Stichprobenumfang (d. h. Anzahl gezogener Hausumringe) in Schicht  $h$  ( $h = 1, 2, \dots, 5$ ) des Erhebungsbezirks  $i$
- $N_i(h)$  : Anzahl der sekundären Erhebungseinheiten (d. h. Hausumringe) in Schicht  $h$  ( $h = 1, 2, \dots, 5$ ) des Erhebungsbezirks  $i$

Zu beachten ist in diesem Kontext, dass die Anzahl der Nichtwohngebäude in Erhebungsbezirk  $i$  – mengentheoretisch gesprochen die Mächtigkeit von  $U_i$  – unbekannt ist und nur im Falle eines (unterbliebenen) vollständigen Screenings von Erhebungsbezirk  $i$  ermittelbar gewesen wäre.

In **Fallkonstellation 2**, in der ein Hausumring Schnittflächen mit mehr als einem Nichtwohngebäude hat, Letztere aber nicht von weiteren Hausumringen geschnitten werden, hängt  $\pi_{k|i}$  davon ab, wie viele dieser Gebäude – alle oder nur ein Teil davon – in die Untersuchung aufgenommen und gescreent werden. Die Entscheidung fiel aus den folgenden erhebungspraktischen und stichprobenmethodischen Gründen für eine Vollerhebung aller einem Stichproben-Hausumring zugeordneten Nichtwohngebäude:

- Im Screening sollen auf der Grundlage der gezogenen Hausumring-Stichprobe so viele Nichtwohngebäude wie möglich in die Untersuchung aufgenommen werden (vgl. Unterkapitel 3.2.1), wodurch nicht einzusehen ist, warum auf die Erhebung aller mit dem Stichproben-Hausumring in Verbindung stehenden Nichtwohngebäude verzichtet werden sollte.
- Der Zusatzaufwand infolge der Erhebung von Daten zu allen zum Stichproben-Hausumring „gehörenden“ Nichtwohngebäuden ist vertretbar gering, zumal das Screening-Personal zwecks Aufsuchens des Stichproben-Hausumrings ohnehin schon vor Ort ist.
- Die beim Verzicht auf eine Vollerhebung obligatorische Ziehung einer Zufallsstichprobe<sup>56</sup> müsste vor Ort durch das Screening-Personal erfolgen, was Zusatzaufwand hervorrufen würde und fehleranfällig wäre.
- Im Falle einer Zufallsstichprobe käme eine dritte Ziehungsstufe hinzu, was die stichprobenmethodische Modellierung und hierbei insbesondere die Ableitung von Ziehungswahrscheinlichkeiten 1. und 2. Ordnung weiter verkomplizieren würde.

Aufgrund der Entscheidung zugunsten einer Vollerhebung entsprechen in Fallkonstellation 2 die Ziehungswahrscheinlichkeiten aller einem Stichproben-Hausumring zugeordneten Nichtwohngebäude der Ziehungswahrscheinlichkeit des Stichproben-Hausumrings. Die gebäudebezogenen Ziehungswahrscheinlichkeiten bestimmten sich daher wie auch in Fallkonstellation 1 nach Formel (3).

In **Fallkonstellation 3** zeigen mehrere Erhebungseinheiten auf dieselbe Untersuchungseinheit, d. h. ein Nichtwohngebäude hat Schnittflächen mit mehreren Hausumringen, wobei Letztere nicht in Verbindung mit weiteren Nichtwohngebäuden stehen. Nichtwohngebäude, für die Fallkonstellation 3 gilt, haben dadurch

<sup>56</sup> Eine solche Zufallsstichprobe müsste einen Stichprobenumfang von mindestens zwei haben, da anderweitig die in Kapitel 2 aufgeführte Bedingung, wonach sicherzustellen ist, dass zwei beliebige Nichtwohngebäude eine positive gemeinsame Ziehungswahrscheinlichkeit aufweisen, nicht zu erfüllen wäre. Faktisch käme somit nur dann eine Stichprobenziehung infrage, wenn dem Stichproben-Hausumring mindestens drei Nichtwohngebäude zugeordnet sind.



mehrere Möglichkeiten, in die Untersuchung zu gelangen und gescreent zu werden, als Nichtwohngebäude, die nur über einen einzigen Hausumring erreicht werden können (Fallkonstellationen 1 und 2). Entsprechend sind ihre Ziehungswahrscheinlichkeiten höher. Die Bestimmung dieser Ziehungswahrscheinlichkeiten erfolgt in Anlehnung an (Lessler, Kalsbeek 1992 S. 95), die – wie auch das vorliegende Forschungsvorhaben – von Ziehen ohne Zurücklegen ausgehen<sup>57</sup>. Zunächst sei davon ausgegangen, dass die auf das Nichtwohngebäude  $k$  zeigenden Hausumringe derselben (Relevanzwahrscheinlichkeits-) Schicht  $h$  angehören. Für das Nichtwohngebäude  $k$  hat man daher folgende (bedingte) Ziehungswahrscheinlichkeit:

$$(4) \quad \pi_{k|i} = 1 - \frac{\binom{N_i(h) - M_{k|i}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}} \text{ für } k \in U_i \text{ mit}$$

$M_{k|i}(h)$ : Zahl der auf die Untersuchungseinheit  $k$  (d. h. das Nichtwohngebäude  $k$ ) zeigenden sekundären Erhebungseinheiten (d. h. Hausumringe) der Schicht  $h$  ( $h = 1, 2, \dots, 5$ ) des Erhebungsbezirks  $i$

Der Quotient in Gleichung (4) setzt die Zahl aller wiederholungsfreien<sup>58</sup> schichtenspezifischen Hausumring-Stichproben, in denen die Untersuchungseinheit  $k$ , also das Nichtwohngebäude  $k$ , nicht enthalten ist (Zähler)<sup>59</sup>, ins Verhältnis zur Gesamtzahl aller wiederholungsfreien schichtenspezifischen Hausumring-Stichproben (Nenner). Da alle wiederholungsfreien schichtenspezifischen Hausumring-Stichproben gleichwahrscheinlich sind, gibt der Quotient somit die Wahrscheinlichkeit dafür an, dass die Untersuchungseinheit  $k$ , also das Nichtwohngebäude  $k$ , nicht in die schichtenspezifische Nichtwohngebäudestichprobe gelangt und somit nicht gescreent wird. Dies ist genau dann der Fall, wenn kein Hausumring gezogen wird, der auf das Nichtwohngebäude  $k$  zeigt. Indem der Quotient in Gleichung (4) von der Zahl 1 subtrahiert wird, erhält man die gesuchte Ziehungswahrscheinlichkeit für das Nichtwohngebäude  $k$ .

Die Binomialkoeffizienten in Zähler und Nenner des in Gleichung (4) enthaltenen Quotienten können unvorstellbar groß sein. Allerdings lässt sich Gleichung (4) in einen „berechenbareren“ Ausdruck umformen:

$$(5) \quad \pi_{k|i} = 1 - \prod_{j=0}^{M_{k|i}(h)-1} \frac{N_i(h) - n_i(h) - j}{N_i(h) - j} \text{ für } M_{k|i}(h) \geq 1,$$

<sup>57</sup> Aufgrund des Ziehens ohne Zurücklegen haben Nichtwohngebäude, auf die  $x$  Hausumringe zeigen, nicht eine  $x$ -fach so hohe (bedingte) Ziehungswahrscheinlichkeit wie Nichtwohngebäude, die nur über einen einzigen Hausumring erreicht werden können.

<sup>58</sup> Die Wiederholungsfreiheit ist ein Reflex des Ziehens ohne Zurücklegen und bedeutet, dass ein und derselbe Hausumring nicht mehrmals gezogen werden kann.

<sup>59</sup> Da sich die Stichprobenziehung auf der zweiten Ziehungsstufe auf Hausumringe und nicht auf Nichtwohngebäude erstreckt, können Nichtwohngebäude streng genommen nicht Bestandteil der Hausumring-Stichprobe sein. Sprachlich korrekt müsste es deshalb heißen: „Der Quotient in Gleichung (4) setzt die Zahl aller wiederholungsfreien schichtenspezifischen Hausumring-Stichproben, in denen Hausumringe, die auf die Untersuchungseinheit  $k$ , d. h. das Nichtwohngebäude  $k$ , zeigen, nicht enthalten sind (Zähler), ...“. Da jedoch stets alle Nichtwohngebäude, auf die ein in die Stichprobe aufgenommener Hausumring zeigt, im Screening erhoben werden, wird hier und im Folgenden aus Gründen der sprachlichen Vereinfachung die sprachlich korrekte Formulierung nicht verwendet.

wobei man für  $M_{k|i}(h) = 1$  (Fallkonstellationen 1 und 2) zu Gleichung (3) gelangt<sup>60</sup>.

Angesichts des Wertebereichs, in den  $M_{k|i}(h)$  typischerweise fällt (im Regelfall handelt es sich um eine kleine einstellige Zahl), entstehen somit mühelos berechenbare Ausdrücke.

Wie erwähnt, geht Gleichung (4) davon aus, dass alle fraglichen Erhebungseinheiten derselben Schicht  $h$  angehören. Stammen die auf die Untersuchungseinheit  $k$  zeigenden Erhebungseinheiten, d. h. die mit dem Nichtwohngebäude  $k$  Schnittflächen bildenden Hausumringe, dagegen aus unterschiedlichen Schichten (desselben Erhebungsbezirks  $i$ ), hat man stattdessen

$$(6) \quad \pi_{k|i} = 1 - \prod_{h=1}^5 \frac{\binom{N_i(h)-M_{k|i}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}} \text{ für } k \in U_i.$$

Das Produkt der Quotienten in Gleichung (6) gibt dabei die Wahrscheinlichkeit für das Ereignis an, dass in keiner Schicht Erhebungseinheiten gezogen werden, die auf die Untersuchungseinheit  $k$  zeigen. Das komplementäre Ereignis dazu ist dadurch gekennzeichnet, dass in mindestens einer Schicht mindestens eine Erhebungseinheit gezogen wird, die auf die Untersuchungseinheit  $k$  zeigt. Folglich hat dieses komplementäre Ereignis und damit der Umstand, dass die Untersuchungseinheit  $k$  erhoben wird, eine Wahrscheinlichkeit, die 1 minus dem Produkt der Quotienten in Gleichung (6) entspricht.

Für Schichten, in denen sich überhaupt keine Erhebungseinheiten befinden, die auf die Untersuchungseinheit  $k$  zeigen, ist  $M_{k|i}(h) = 0$ . Für diese Schichten nimmt der Quotient der Binomialkoeffizienten in Gleichung (6) den Wert 1 an. Gehören sogar ausnahmslos alle auf die Untersuchungseinheit  $k$  zeigenden Erhebungseinheiten derselben Schicht an, gelangt man zu Gleichung (4), die daher ein Spezialfall von Gleichung (6) ist. Gleichung (6) stellt somit die grundlegende Formel zur Berechnung untersuchungseinheitsbezogener Inklusionswahrscheinlichkeiten 1. Ordnung dar, die auf alle denkbaren Beziehungsgeflechte zwischen Hausumringen und Nichtwohngebäuden anwendbar ist<sup>61</sup>.

Aufgrund der Möglichkeit, dass auf die Untersuchungseinheit  $k$  Erhebungseinheiten aus unterschiedlichen Schichten zeigen, wäre es übrigens unangemessen, den gebäudebezogenen Ziehungswahrscheinlichkeiten eine Schichtenbezeichnung hinzuzufügen, also beispielsweise  $\pi_{k|i}(h)$  zu schreiben. Die Schichtenbildung auf der zweiten Ziehungsstufe erstreckt sich nur auf die Hausumringe als (sekundäre) Erhebungseinheiten und nicht auf die eigentlich interessierenden Nichtwohngebäude.

Ähnlich wie Gleichung (4) kann auch Gleichung (6) in einen leichter berechenbaren Ausdruck umgeformt werden, nämlich in

$$(7) \quad \pi_{k|i} = 1 - \prod_{h \in H} \prod_{j=0}^{M_{k|i}(h)-1} \frac{N_i(h)-n_i(h)-j}{N_i(h)-j} \text{ für } M_{k|i}(h) \geq 1.$$

<sup>60</sup> Insofern ist Gleichung (3) ein Spezialfall von Gleichung (4).

<sup>61</sup> Wie in Kürze gezeigt wird, gilt Gleichung (6) auch für Fallkonstellation 4.

Während es möglich ist, dass Erhebungseinheiten aus unterschiedlichen Schichten desselben Erhebungsbezirks  $i$  auf ein und dieselbe Untersuchungseinheit  $k$  zeigen, ist das Szenario von auf dieselbe Untersuchungseinheit  $k$  zeigenden Erhebungseinheiten aus unterschiedlichen Erhebungsbezirken ausgeschlossen. Denn bei der Bildung der Erhebungsbezirke wurde darauf geachtet, dass die Erhebungsbezirksgrenzen entlang von Barrieren verlaufen, die aus rechtlichen oder natürlichen Gründen nicht bebaubar sind (z. B. Straßen, Flüsse), wodurch es ausgeschlossen ist, dass Erhebungsbezirksgrenzen zusammenhängende Bebauungen trennen.<sup>62</sup>

Da **Fallkonstellation 4** eine Kombination aus den beiden Fallkonstellationen 2 und 3 ist, lassen sich die Gebäudeziehungswahrscheinlichkeiten bei Vorliegen dieser Fallkonstellation mit Verweis auf die bisherigen Schilderungen bestimmen:

- Für Gebäude, auf die ausschließlich ein einziger Hausumring zeigt, entspricht die Ziehungswahrscheinlichkeit exakt der Ziehungswahrscheinlichkeit dieses Hausumrings.
- Für die Ziehungswahrscheinlichkeit von Gebäuden, auf die mehrere Hausumringe zeigen, ist Gleichung (4) oder (6) (bzw. die leichter zu berechnenden Ausdrücke (5) oder (7)) maßgebend.

Damit untersuchungseinheitsbezogene, d. h. nichtwohngebäudespezifische Ziehungswahrscheinlichkeiten mithilfe der beschriebenen Formeln berechnet werden können, muss für jedes Nichtwohngebäude  $k$ , auf das ein Stichproben-Hausumring zeigt,  $M_{k|i}(h)$  bekannt sein. Aus erhebungspraktischen Gründen hat das Screening-Personal jedoch nicht direkt  $M_{k|i}(h)$  bestimmt, sondern zu jedem Nichtwohngebäude<sup>63</sup>, dessen Grundfläche über den aufzusuchenden Hausumring hinaus durch weitere Hausumringe überdeckt wird, die Identifikationsnummern dieser Hausumringe vermerkt. Die Bestimmung von  $M_{k|i}(h)$  erfolgte dann auf Basis dieser Angaben an zentraler Stelle. Damit das Screening-Personal die benötigten Angaben machen konnte, war zum einen eine entsprechende Unterweisung erforderlich. Zum anderen musste gewährleistet werden, dass aus einem Lageplan nicht nur der aufzusuchende Stichproben-Hausumring hervorgeht, sondern darüber hinaus auch die benachbarten Hausumringe mitsamt Identifikationsnummern – unabhängig davon, ob diese Hausumringe ebenfalls in die Hausumring-Stichprobe Eingang gefunden haben.

## Inklusionswahrscheinlichkeiten 2. Ordnung

Die untersuchungseinheitsbezogenen Inklusionswahrscheinlichkeiten 2. Ordnung geben im vorliegenden Fall die Wahrscheinlichkeit dafür an, dass zwei beliebige Nichtwohngebäude gezogen und daher im Rahmen des Screenings zum Zwecke der Datenaufnahme aufgesucht werden. Dies ist dann der Fall, wenn für jedes der beiden Nichtwohngebäude mindestens ein Hausumring in die Hausumring-Stichprobe gelangt, der auf dieses Nichtwohngebäude zeigt.

<sup>62</sup> Anders ausgedrückt, kann jeder Erhebungsbezirk entlang seiner Grenze abgeschritten werden, ohne dass eine Gebäudewand den Weg versperren würde. Zur Bildung der Erhebungsbezirke vgl. Unterkapitel 3.2.2.

<sup>63</sup> Es sei an dieser Stelle daran erinnert, dass alle Nichtwohngebäude erhoben wurden, mit denen der aufzusuchende Stichproben-Hausumring Schnittflächen aufwies.



Die beiden betrachteten Untersuchungseinheiten bzw. Nichtwohngebäude, die im Folgenden mit  $k$  und  $l$  (wobei  $k \neq l$ ) bezeichnet werden, können, müssen dabei nicht demselben Erhebungsbezirk angehören. Davon abhängig, gilt für  $\pi_{kl}$  als der gesuchten Inklusionswahrscheinlichkeiten 2. Ordnung bezüglich der beiden Nichtwohngebäude  $k$  und  $l$ :

$$(8) \quad \pi_{kl} = \begin{cases} \pi_{li}\pi_{kl|i} & \text{für } k, l \in U_i \\ \pi_{lij}\pi_{k|i}\pi_{l|j} & \text{für } k \in U_i \text{ und } l \in U_j \text{ mit } i \neq j \end{cases}$$

Während die beiden Nichtwohngebäude im ersten Fall (obere Zeile von Gleichung (8)) im selben Erhebungsbezirk  $i$  liegen, gehören sie im zweiten Fall (untere Zeile von Gleichung (8)) unterschiedlichen Erhebungsbezirken, nämlich  $i$  und  $j$ , an.

Zu spezifizieren sind noch  $\pi_{kl|i}$  (obere Zeile von Gleichung (8)) und  $\pi_{lij}$  (untere Zeile von Gleichung (8)).

$\pi_{kl|i}$  ist die bedingte Inklusionswahrscheinlichkeit 2. Ordnung bezüglich der beiden Nichtwohngebäude  $k$  und  $l$  unter der Bedingung der vorgelagerten Ziehung ihres gemeinsamen Erhebungsbezirks  $i$ .  $\pi_{kl|i}$  lässt sich nach folgender, sich ggf. noch stark vereinfachender Gleichung berechnen:

$$(9) \quad \pi_{kl|i} = 1 - A_{kl|i} \text{ mit}$$

$$A_{kl|i} = \prod_{h=1}^5 \frac{\binom{N_i(h) - M_{kl|i}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}} + \prod_{h=1}^5 \frac{\binom{N_i(h) - M_{li}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}} - \prod_{h=1}^5 \frac{\binom{N_i(h) - M_{kl|i}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}}$$

Dabei ist

$M_{kl|i}(h)$ : Zahl der sekundären Erhebungseinheiten (d. h. Hausumringe) der Schicht  $h$  ( $h = 1, 2, \dots, 5$ ), die auf die Untersuchungseinheit  $k$  (d. h. das Nichtwohngebäude  $k$ ) und/oder die Untersuchungseinheit  $l$  (d. h. das Nichtwohngebäude  $l$ ) zeigen<sup>64</sup>

Gleichung (9) ist eine Anwendung zweier Rechenregeln der Wahrscheinlichkeitstheorie, zum einen des Additionssatzes der Wahrscheinlichkeitstheorie und zum anderen der Formel zur Berechnung der Wahrscheinlichkeiten für komplementäre Ereignisse. Konkret ist

$$(10) \quad W(A \cap B) = 1 - W(\overline{A \cap B}) = 1 - [W(\overline{A}) + W(\overline{B}) - W(\overline{A \cup B})],$$

wobei das Ereignis  $A$  ( $B$ ) darauf abstellt, ob Untersuchungseinheit  $k$  ( $l$ ) in die Stichprobe aufgenommen wird.

Gleichung (9) macht keinerlei Vorgaben, aus welchen Schichten die Erhebungseinheiten stammen, die auf die Untersuchungseinheit  $k$  bzw.  $l$  zeigen, und wie viele Erhebungseinheiten es jeweils sind. Vielmehr kann auf jede der beiden

<sup>64</sup> Zeigt eine Erhebungseinheit auf beide Untersuchungseinheiten, wird sie nur einmal gezählt.

Untersuchungseinheiten mehr als eine Erhebungseinheit zeigen und diese Erhebungseinheiten können unterschiedlichen Schichten angehören.

Gehören alle Erhebungseinheiten, die auf die Untersuchungseinheiten  $k$  bzw.  $l$  zeigen, derselben Schicht  $h$  an, vereinfacht sich Gleichung (9) wie folgt:

$$(11) \quad \pi_{kl|i} = 1 - \frac{\binom{N_i(h)-M_{k|i}(h)}{n_i(h)} + \binom{N_i(h)-M_{l|i}(h)}{n_i(h)} - \binom{N_i(h)-M_{kl|i}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}}$$

Der erste Summand im Zähler des Quotienten von Gleichung (11) entspricht der Zahl aller wiederholungsfreien schichtenspezifischen Stichproben, in denen die Untersuchungseinheit  $k$  nicht enthalten ist. Vergleichbares gilt für den zweiten Summanden, der auf die andere Untersuchungseinheit abstellt. Da in einem Teil der Stichproben, in denen die eine Untersuchungseinheit (z. B. die Untersuchungseinheit  $k$ ) nicht enthalten ist, die andere Untersuchungseinheit (im Beispiel die Untersuchungseinheit  $l$ ) ebenfalls nicht aufgenommen wurde, ist zur Vermeidung von Doppelzählungen von der Summe der ersten beiden Summanden im Zähler des Quotienten von Gleichung (11) die Zahl derjenigen Stichproben abziehen, in denen weder die eine noch die andere Untersuchungseinheit enthalten ist. Der Zähler des Quotienten in Gleichung (11) gibt somit die Zahl aller wiederholungsfreien schichtenspezifischen Stichproben an, in denen mindestens eine der beiden Untersuchungseinheiten fehlt. Der Quotient selbst stellt dagegen auf den Anteil dieser Stichproben an allen wiederholungsfreien schichtenspezifischen Stichproben ab und gibt aufgrund des Umstandes, dass all diese Stichproben gleichwahrscheinlich sind, die Wahrscheinlichkeit dafür an, dass in der gezogenen Stichprobe wenigstens eine der beiden Untersuchungseinheiten fehlt. Indem man diese Wahrscheinlichkeit von der Zahl 1 abzieht, erhält man die gesuchte Inklusionswahrscheinlichkeit 2. Ordnung.

Weist auf jede der beiden Untersuchungseinheiten  $k$  und  $l$  nur je eine, aber nicht dieselbe Erhebungseinheit, d. h. haben die beiden Nichtwohngebäude mit jeweils nur einem einzigen, aber nicht mit demselben Hausumring Schnittflächen, und gehören überdies beide Erhebungseinheiten derselben Schicht  $h$  an, vereinfacht sich Gleichung (11) wegen  $M_{k|i}(h) = M_{l|i}(h) = 1$  und  $M_{kl|i}(h) = M_{k|i}(h) + M_{l|i}(h) = 2$  weiter zu

$$(12) \quad \pi_{kl|i} = \frac{n_i(h)[n_i(h)-1]}{N_i(h)[N_i(h)-1]}.$$

Liegt dagegen Fallkonstellation 2 aus Abbildung 3-2 vor, d. h. zeigt auf beide Untersuchungseinheiten  $k$  und  $l$  ausschließlich eine und überdies dieselbe Erhebungseinheit, hat man wegen  $M_{k|i}(h) = M_{l|i}(h) = M_{kl|i}(h) = 1$

$$(13) \quad \pi_{kl|i} = \pi_{k|i} = \pi_{l|i} = \frac{n_i(h)}{N_i(h)}.$$

$\pi_{ij}$  in der unteren Zeile von Gleichung (8) steht für die Inklusionswahrscheinlichkeit 2. Ordnung bezüglich der Erhebungsbezirke  $i$  und  $j$ , also für die Wahrscheinlichkeit, dass auf der ersten Ziehungsstufe sowohl der Erhebungsbezirk  $i$  als auch der Erhebungsbezirk  $j$  gezogen werden. Je nachdem, ob die beiden Erhebungsbezirke derselben Schicht angehören, hat man für  $\pi_{ij}$

$$(14) \quad \pi_{ij} = \begin{cases} \frac{n_I(g)[n_I(g)-1]}{N_I(g)[N_I(g)-1]} & \text{für } i, j \in U_{Ig} \\ \frac{n_I(g)}{N_I(g)} \frac{n_I(q)}{N_I(q)} & \text{für } i \in U_{Ig} \text{ und } j \in U_{Iq} \text{ mit } g \neq q \end{cases}$$

### 3.3 Schätzung

#### 3.3.1 Schätzung von Merkmalssummen

Unter einer Merkmalssumme in der Grundgesamtheit versteht man die Summe aller Merkmalsausprägungen eines interessierenden Merkmals  $y$  über alle Untersuchungseinheiten. Mathematisch lässt sich eine Merkmalssumme  $t$  ausdrücken als

$$(15) \quad t = \sum_U y_k \text{ mit}$$

$U$  : Menge aller Untersuchungseinheiten (d. h. aller Nichtwohngebäude)

$y_k$  : Merkmalsausprägung des interessierenden Merkmals  $y$  von Untersuchungseinheit  $k$  (d. h. von Nichtwohngebäude  $k$ )

Zu beachten ist, dass es hinsichtlich der Skalierung des interessierenden Merkmals  $y$  keinerlei Einschränkungen gibt. Es kann sich unter anderem auch um ein dichotomes Merkmal mit den beiden „künstlich gesetzten“ Merkmalsausprägungen „0“ und „1“ und damit um eine Indikatorvariable handeln, die auf das Vorliegen einer bestimmten Eigenschaft abstellt (z. B. mit Fernwärme beheizt), oder aber um ein metrisches Merkmal (z. B. Nettoraumfläche). Beispiele für den Populationsparameter  $t$  sind daher die Anzahl der Nichtwohngebäude in Deutschland, die mit Fernwärme beheizt werden, und die gesamte Nettoraumfläche des Nichtwohngebäudebestands.

Um den Populationsparameter  $t$  auf Stichprobenbasis zu schätzen, greift man gemeinhin und so auch im vorliegenden Projekt auf den sog. Horvitz-Thompson- (Merkmalssummen-) Schätzer (kurz: HT-Schätzer)  $\hat{t}_\pi$  zurück, für den

$$(16) \quad \hat{t}_\pi = \sum_U L_k \frac{y_k}{\pi_k} \quad \forall \pi_k > 0 \text{ und } k \in U \text{ mit}$$

$L_k$  : Indikatorvariable, die auf die Zugehörigkeit von Untersuchungseinheit  $k$  (d. h. von Nichtwohngebäude  $k$ ) zur gezogenen Stichprobe<sup>65</sup> abstellt

gilt.

Da die Indikatorvariable für nicht in die Stichprobe aufgenommene Nichtwohngebäude den Wert 0 annimmt, gehen in die Berechnung von (16) faktisch nur die Merkmalsausprägungen und Ziehungswahrscheinlichkeiten derjenigen Nichtwohngebäude ein, die der gezogenen Stichprobe angehören, also im Screening identifiziert wurden, so dass die „eigentliche“ Berechnung nach der Formel

$$(17) \quad \hat{t}_\pi = \sum_s \frac{y_k}{\pi_k} \quad \forall \pi_k > 0 \text{ und } k \in s \text{ mit}$$

<sup>65</sup> Gemeint sind diejenigen Nichtwohngebäude, die mit Hausumringen Schnittflächen aufweisen, die auf der zweiten Ziehungsstufe ausgewählt wurden.

$s$  : Menge der Untersuchungseinheiten (d. h. also der Nichtwohngebäude), die über die gezogene Stichprobe identifiziert worden sind

erfolgt.

Dadurch, dass der HT-Schätzer die Merkmalsausprägung jedes Stichproben-Nichtwohngebäudes durch die korrespondierende Gebäudeziehungswahrscheinlichkeit dividiert und danach über alle Stichproben-Nichtwohngebäude aufsummiert, werden die Merkmalsausprägungen von Stichproben-Nichtwohngebäuden mit hoher Ziehungswahrscheinlichkeit gering gewichtet und umgekehrt die mit niedriger Ziehungswahrscheinlichkeit entsprechend hoch. Der Kehrwert der Ziehungswahrscheinlichkeit eines Nichtwohngebäudes entspricht dabei dessen Hochrechnungsfaktor, der angibt, für wie viele Nichtwohngebäude der Grundgesamtheit das Stichproben-Nichtwohngebäude stellvertretend steht (vgl. auch Fußnote 48).

Mithilfe von (16) bzw. (17) kann auch die Gesamtzahl aller Nichtwohngebäude in Deutschland geschätzt werden, denn diese Gesamtzahl entspricht ebenfalls einer Merkmalssumme, nämlich derjenigen, bei der das Merkmal  $y$  für alle Stichproben-Nichtwohngebäude die Ausprägung 1 annimmt. Der HT-Schätzer für den Umfang der Grundgesamtheit, also für die Gesamtzahl aller Nichtwohngebäude in Deutschland, entspricht daher den aufsummierten Hochrechnungsfaktoren aller im Screening identifizierten Nichtwohngebäude. Zu beachten ist in diesem Kontext, dass der Umfang der Grundgesamtheit im vorliegenden Projektkontext nicht ohne Unsicherheit beziffert, sondern nur aus der erhobenen Stichprobe heraus geschätzt werden kann. Dafür zeichnen zwei Defizite der Auswahlgrundlage verantwortlich, zum einen die Existenz irrelevanter Hausumringe in der Auswahlgrundlage (vgl. Unterkapitel 3.1.3) und zum anderen die nicht notwendigerweise gegebene eindeutige Beziehung zwischen Hausumring und Nichtwohngebäude (vgl. Unterkapitel 3.1.2).

Es kann gezeigt werden, dass  $\hat{t}_\pi$  unter der hier erfüllten Voraussetzung positiver Ziehungswahrscheinlichkeiten ein erwartungstreu bzw. unverzerrter Schätzer für  $t$  ist (vgl. z. B. (Särndal et al. 1992 S. 42)). Das bedeutet, dass die mit  $\hat{t}_\pi$  ermittelten Schätzwerte im Mittel über alle möglichen Stichproben, die nach demselben Auswahlverfahren hätten gezogen werden können,  $t$  exakt treffen, sofern bei dieser Mittelung jede dieser möglichen Stichproben mit der Wahrscheinlichkeit ihres Eintretens gewichtet wird (vgl. auch Kapitel 2). Das Vorhandensein irrelevanter Einheiten wie z. B. im Screening identifizierte Wohngebäude wirkt sich dabei nicht störend auf die Erwartungstreue aus: Diese bleiben bei der Schätzung einfach unberücksichtigt.

Die gängige Maßzahl für die Stärke, mit der  $\hat{t}_\pi$  von Stichprobe zu Stichprobe schwankt und damit um  $t$  streut, ist die sog. Varianz (*variance*), für die folgende Beziehung gilt (vgl. ebenda, S. 43):

$$(18) \quad V(\hat{t}_\pi) = \sum \sum_U (\pi_{kl} - \pi_k \pi_l) \frac{y_k y_l}{\pi_k \pi_l} \quad \forall \pi_k, \pi_l > 0 \text{ und } k, l \in U$$

Zieht man die Wurzel aus  $V(\hat{t}_\pi)$ , erhält man den Standardfehler – die gängige Kennziffer zur Beschreibung der stichprobenimmanenten Schätzungenauigkeit.

Ob die gezogene Stichprobe nun einen Schätzwert liefert, der mit dem tatsächlichen Wert von  $t$  übereinstimmt bzw. wie weit andernfalls dieser Schätzwert von  $t$  abweicht, kann ohne Kenntnis von  $t$  naturgemäß nicht angegeben werden. Aber

auch  $V(\hat{t}_\pi)$  als Maß für die Stärke der Streuung von  $\hat{t}_\pi$  lässt sich auf Basis von Stichprobendaten nicht berechnen (vgl. Aufsummierung über  $U$ ). Mittels der gezogenen Stichprobe kann unter der zusätzlichen und hier erfüllten Annahme positiver Inklusionswahrscheinlichkeiten 2. Ordnung jedoch ein erwartungstreuer Schätzer für  $V(\hat{t}_\pi)$  berechnet werden, nämlich

$$(19) \quad \hat{V}(\hat{t}_\pi) = \sum_s \sum_{kl} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k y_l}{\pi_k \pi_l} \quad \forall \pi_k, \pi_l, \pi_{kl} > 0 \text{ und } k, l \in s$$

(vgl. ebenda). Indem man die Wurzel aus  $\hat{V}(\hat{t}_\pi)$  zieht, gelangt man zu einem (approximativ) erwartungstreuen Schätzer für den Standardfehler. Der aus der gezogenen Stichprobe abgeleitete Schätzwert für den Standardfehler wird in diesem Projekt zusammen mit dem geschätzten Populationsparameter stets ausgewiesen.

Die bisherigen Ausführungen zur Merkmalssummenschätzung unterstellten, dass datenorientierte Aussagen über die Grundgesamtheit aller Nichtwohngebäude in Deutschland abgeleitet werden sollen. Folgerichtig wurde in den Formeln (15), (16) und (18) über alle Nichtwohngebäude ( $U$ ) und in den Formeln (17) und (19) über alle im Screening identifizierten Nichtwohngebäude ( $s$ ) summiert. Sollen dagegen nur Auswertungen für eine Teilmenge  $\tilde{U}$  der Grundgesamtheit ( $\tilde{U} \subset U$ ), also Auswertungen für ein bestimmtes Gebäudesegment (z. B. GEG-relevante Nichtwohngebäude, Schulen, Nichtwohngebäude in Ostdeutschland) vorgenommen werden, gelten die Formeln entsprechend. Dazu ist lediglich  $U$  durch  $\tilde{U}$  bzw.  $s$  durch  $\tilde{s}$  zu ersetzen, wobei die Tilde kennzeichnet, das jeweils nur ein bestimmtes Gebäudesegment, also eine bestimmte Subpopulation (*domain*) der Grundgesamtheit, interessiert. Zu beachten ist in diesem Kontext, dass – wie erwähnt – nicht nur die Gesamtzahl aller Nichtwohngebäude in Deutschland aus der gezogenen Stichprobe heraus geschätzt werden muss, sondern auch der Besetzungsumfang jeder Teilmenge davon. Aufgrund dessen kann nicht nur die Gesamtzahl aller Nichtwohngebäude in Deutschland, sondern z. B. auch die Gesamtzahl aller GEG-relevanten Nichtwohngebäude nicht exakt beziffert, sondern nur geschätzt werden – verbunden mit der damit einhergehenden Unsicherheit, die sich in positiven Standardfehlern niederschlägt. Dieser Umstand hat nicht zuletzt Auswirkungen auf die Schätzung des Verhältnisses zweier Merkmalssummen, die im folgenden Unterkapitel behandelt wird.

### 3.3.2 Schätzung des Verhältnisses zweier Merkmalssummen

Beim Verhältnis (*ratio*) zweier Merkmalssummen in der Grundgesamtheit werden die über alle Untersuchungseinheiten aufsummierten Merkmalsausprägungen eines interessierenden Merkmals  $y$  in Beziehung zu den ebenfalls über alle Untersuchungseinheiten aufsummierten Merkmalsausprägungen eines anderen interessierenden Merkmals  $z$  gesetzt.

Mathematisch lässt sich dieses Verhältnis  $R$  daher ausdrücken als

$$(20) \quad R = \frac{t_y}{t_z} = \frac{\sum_U y_k}{\sum_U z_k} \text{ mit}$$

$z_k$  : Merkmalsausprägung des interessierenden Merkmals  $z$  von Untersuchungseinheit  $k$  (d. h. von Nichtwohngebäude  $k$ )

Formel (20) setzt keinerlei Restriktionen hinsichtlich der Skalierung der Merkmale  $y$  und  $z$ . Insbesondere kann es sich auch um metrische Merkmale handeln. Stellt ENOB:dataNWG

das Merkmal  $y$  beispielsweise auf den Energieverbrauch in kWh eines Nichtwohngebäudes in einem bestimmten Zeitraum ab und steht  $z$  für die Nettonraumfläche in  $m^2$ , entspricht  $R$  dem Energieverbrauch in kWh pro  $m^2$  Nettonraumfläche im deutschen Nichtwohngebäudebestand.

Handelt es sich bei  $z$  um eine Indikatorvariable, die auf die Zugehörigkeit einer Untersuchungseinheit  $k$  (d. h. von Nichtwohngebäude  $k$ ) zur Grundgesamtheit abstellt<sup>66</sup>, und bei  $y$  um eine Indikatorvariable für die Zugehörigkeit zu einer interessierenden Teilpopulation der Grundgesamtheit, entspricht  $R$  dem Anteil, den diese Teilpopulation am gesamten Nichtwohngebäudebestand einnimmt.

Ein mit Stichprobendaten berechenbarer und im vorliegenden Projekt verwendeter Verhältnisschätzer (*ratio estimator*)  $\hat{R}$  für  $R$  ist der Quotient zweier HT-Merkmalssummenschatzer  $\hat{t}_{y\pi}$  und  $\hat{t}_{z\pi}$ :

$$(21) \quad \hat{R} = \frac{\hat{t}_{y\pi}}{\hat{t}_{z\pi}} = \frac{\sum_s y_k / \pi_k}{\sum_s z_k / \pi_k}$$

Obwohl im Zähler und Nenner von  $\hat{R}$  jeweils erwartungstreue Schätzer stehen, ist  $\hat{R}$  nur näherungsweise bzw. approximativ erwartungstreu (vgl. (Särndal et al. 1992 S. 176/177).

Die (approximative) Varianz von  $\hat{R}$  ist gegeben durch

$$(22) \quad AV(\hat{R}) = \frac{1}{t_z^2} \sum \sum_U (\pi_{kl} - \pi_k \pi_l) \frac{y_k - R z_k}{\pi_k} \frac{y_l - R z_l}{\pi_l} \quad \forall \pi_k, \pi_l > 0 \text{ und } k, l \in U$$

und kann auf Stichprobenbasis nicht berechnet werden (vgl. ebenda, S. 178). Als Varianzschätzer kam bei den Auswertungen daher

$$(23) \quad \hat{V}(\hat{R}) = \frac{1}{\hat{t}_{z\pi}^2} [\hat{V}(\hat{t}_{y\pi}) + \hat{R}^2 \hat{V}(\hat{t}_{z\pi}) - 2\hat{R} \hat{C}(\hat{t}_{y\pi}, \hat{t}_{z\pi})] \quad \forall \pi_k, \pi_l, \pi_{kl} > 0 \text{ und } k, l \in U$$

mit

$$(24) \quad \hat{C}(\hat{t}_{y\pi}, \hat{t}_{z\pi}) = \sum \sum_U L_{kl} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k} \frac{z_l}{\pi_l} = \sum \sum_S \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k} \frac{z_l}{\pi_l} \text{ mit}$$

$L_{kl}$  : Indikatorvariable, die auf die Zugehörigkeit der Untersuchungseinheiten  $k$  und  $l$  (d. h. der Nichtwohngebäude  $k$  und  $l$ ) zur gezogenen Stichprobe abstellt

zum Einsatz (vgl. ebenda, S. 179).  $\hat{C}(\hat{t}_{y\pi}, \hat{t}_{z\pi})$  ist ein erwartungstreuer Kovarianzschätzer für

$$(25) \quad C(\hat{t}_{y\pi}, \hat{t}_{z\pi}) = \sum \sum_U (\pi_{kl} - \pi_k \pi_l) \frac{y_k}{\pi_k} \frac{z_l}{\pi_l},$$

der Kovarianz der beiden HT-Merkmalssummenschatzer  $\hat{t}_{y\pi}$  und  $\hat{t}_{z\pi}$  (vgl. ebenda, S. 170).

Die Ausführungen zur Schätzung des Verhältnisses zweier Merkmalssummen gelten analog auch bei einer gebäudesegmentspezifischen Betrachtung, wobei dann

<sup>66</sup> Im Nenner von Formel (20) steht in diesem Fall der Umfang der Grundgesamtheit, also die Anzahl aller Nichtwohngebäude.

– wie bei der Schätzung von Merkmalssummen (vgl. Unterkapitel 3.3.1) –  $U$  durch  $\bar{U}$  und  $s$  durch  $\bar{s}$  zu ersetzen ist.

### 3.3.3 Kombinierte Schätzung von Merkmalssummen und Gaußsches Fehlerfortpflanzungsgesetz

Die in Unterkapitel 3.3.1 beschriebene Vorgehensweise zur Schätzung von Merkmalssummen ist nur möglich, wenn das betrachtete Merkmal frei von sog. fehlenden Werten (*missing values*) ist. Im Falle einer Schätzung für eine Subpopulation (*domain*) darf darüber hinaus auch dasjenige Merkmal keine fehlenden Werte aufweisen, über das die Subpopulation definiert wird. Die beschriebenen Voraussetzungen sind in den meisten Fällen jedoch nicht gegeben. Aufgrund dessen muss ein anderer Weg zur Schätzung von Merkmalssummen beschritten werden. Bevor dieser Weg erläutert wird, soll zunächst kurz erläutert werden, was sich hinter dem Phänomen fehlender Werte verbirgt und warum es besonderer Beachtung bedarf.

Von fehlenden Werten spricht man, wenn für ein betrachtetes Merkmal nicht für alle Merkmalsträger (hier: Untersuchungsgebäude) gültige Merkmalsausprägungen vorliegen, die Datenmatrix der in der Untersuchung erhobenen Merkmale also entgegen der mit der Datenerhebung verbundenen Intention Lücken aufweist.<sup>67</sup> Die Existenz fehlender Werte ist dabei ein Reflex von Item Nonresponse, das in Unterkapitel 4.1.2 mitsamt der hierfür verantwortlichen Gründe erläutert wird. Auch wenn in der gesamten Untersuchung versucht wurde, Item Nonresponse und damit die Existenz fehlender Werte zu minimieren, gibt es dennoch fast kein Gebäudemerkmal, das hiervon gänzlich verschont blieb. Selbst Merkmale, die nicht auf primärstatistischem Wege in einer der drei Erhebungsphasen erhoben wurden, sondern die den Untersuchungsgebäuden durch die Auswertung bzw. Aufbereitung sekundärstatistischer Datenmaterials nachträglich zugespielt wurden (z.B. Merkmale, die die Gebäudehüllfläche beschreiben), sind nur in Ausnahmefällen frei von fehlenden Werten.<sup>68</sup>

Fehlende Werte führen bei Merkmalssummenschätzungen, wie sie in Unterkapitel 3.3.1 beschrieben wurden, zu einer systematischen Unterschätzung der tatsächlichen Merkmalssumme, da diejenigen Merkmalsträger, für die in Bezug auf das

<sup>67</sup> Lücken in der Datenmatrix können auch intendiert sein, nämlich dann, wenn bei bestimmten Merkmalsträgern die Erhebung eines Merkmals aus inhaltlichen Erwägungen nicht sinnvoll ist und deshalb unterblieb. Solche intendierten Lücken entstehen immer als Folge von Filter- bzw. Sprunganweisungen im Fragebogen. Denn diese Anweisungen führen dazu, dass in Abhängigkeit bestimmter Angaben zu vorausgegangenen Fragen unzutreffende Folgefragen (komplexe) übersprungen werden. Beispielsweise war der in der Breiterehebung verwendete elektronische Fragebogen so programmiert, dass die Kollektorfläche der Solarwärmanlage nur dann erhoben wurde, wenn der Gebäudeansprechpartner die zuvor gestellte Frage nach der Installation einer solchen Anlage auf dem Dach bzw. an der Gebäudefassade bejaht hat. Untersuchungsgebäude ohne solche Anlagen weisen daher bezüglich des Merkmals „Kollektorfläche der Solarwärmanlage“ intentionsgemäß einen fehlenden Wert auf.

<sup>68</sup> Zu diesen Ausnahmefällen gehören insbesondere raumbezogene Merkmale wie beispielsweise der regionalstatistische Gemeindetyp. Da diese Merkmale flächendeckend vorliegen und jedes Untersuchungsgebäude eindeutig geografisch verortet werden kann, konnten sie ausnahmslos jedem Untersuchungsgebäude zugespielt werden. Gleiches gilt für die Zugehörigkeit von Untersuchungsgebäuden zu Gebietskörperschaften (z.B. Bundesland, Kreis).



interessierende Merkmal Werte fehlen, bei der Schätzung unberücksichtigt bleiben müssen.<sup>69</sup>

Wie dennoch die (unverzerrte) Schätzung von Merkmalssummen beim Vorhandensein fehlender Merkmale gelingt, wird im Folgenden anhand eines Beispiels erläutert.

Im Beispiel soll die Brutto-Grundfläche (BGF)<sup>70</sup> aller GEG-relevanten Nichtwohngebäude geschätzt werden. Dabei ergibt sich das Problem, das für einen Teil derjenigen Untersuchungsgebäude, bei denen es sich um GEG-relevante Nichtwohngebäude handelt, die BGF nicht bekannt ist.<sup>71</sup> Würde man diesen Umstand ignorieren und die gesuchte Merkmalssumme wie in Unterkapitel 3.3.1 beschrieben schätzen, wäre eine systematische Unterschätzung die Konsequenz. Als Abhilfe bietet es sich an, die gesuchte Merkmalssumme als Produkt zweier anderer Schätzgrößen zu schätzen, nämlich eines geeigneten Merkmalssummenschätzwerts für die Populationsgröße, der frei von fehlenden Werten ist und deshalb nach dem in Unterkapitel 3.3.1 beschriebenen Verfahren ermittelt werden kann, und einem geeigneten Verhältnisschätzwert.

Was den Merkmalssummenschätzwert angeht, weist Tabelle 3-6 aus, dass es (hochgerechnet) 1,981 Mio. GEG-relevante Nichtwohngebäude in Deutschland gibt.

**Tabelle 3-6: Anzahl Nichtwohngebäude nach Art des Gebäudes**

|                                                              | Anzahl in Tsd. | absoluter Standardfehler in Tsd. |
|--------------------------------------------------------------|----------------|----------------------------------|
| <b>Nichtwohngebäude</b>                                      | <b>21.124</b>  | <b>445</b>                       |
| ... davon thermisch relevante Nichtwohngebäude               | 2.172          | 168                              |
| ... davon GEG-relevante Nichtwohngebäude                     | 1.981          | 152                              |
| ... davon thermisch gering-konditionierte Nichtwohngebäude   | 192            | 47                               |
| ... davon sonstige thermisch konditionierte Nichtwohngebäude | 4.166          | 170                              |
| ... davon thermisch nicht konditionierte Nichtwohngebäude    | 14.786         | 375                              |

Quelle: Eigene Berechnung und Darstellung (finale Auswertung Pos. 01.01.02)

<sup>69</sup> Beschränkt sich die Merkmalssummenschätzung darüber hinaus auf eine Subpopulation und weist dasjenige Merkmal, über das die Subpopulation identifiziert wird, fehlende Werte auf, die sich dahingehend auswirken, dass tatsächlich zur Subpopulation zählende Merkmalsträger dieser nicht zugeordnet werden können, tritt ein weiterer Grund für eine systematische Unterschätzung der tatsächlichen Merkmalssumme hinzu.

<sup>70</sup> Die BGF entspricht der Summe aller Grundflächen aller Grundrissebenen eines Gebäudes und wurde in der vorliegenden Untersuchung vereinfacht als Produkt aus der Gebäudegrundfläche und der Anzahl der Vollgeschosse, differenziert nach oberirdischen und unterirdischen Vollgeschossen, berechnet.

<sup>71</sup> Von den 5.107 GEG-relevanten Nichtwohngebäuden in der Untersuchung wiesen 29 Gebäude aus unterschiedlichen Gründen fehlende Werte in Bezug auf die BGF auf.

Alle Schätzwerte in der mittleren Spalte von Tabelle 3-6 und somit auch der gesuchte Schätzwert für den Bestand an GEG-relevanten Nichtwohngebäuden in Deutschland konnten dabei über das in Unterkapitel 3.3.1 beschriebene Verfahren zur Schätzung von Merkmalssummen ermittelt werden, denn das zugrunde liegende Merkmal, die Gebäudeart, war für ausnahmslos jedes in die Untersuchung aufgenommene Nichtwohngebäude bekannt.

Während der Merkmalssummenschätzwert zur Ermittlung des Populationsumfangs – im Beispiel die Anzahl GEG-relevanter Nichtwohngebäude in Deutschland – zwingend frei von fehlenden Werten sein muss, ist diese Voraussetzung für den Verhältnisschätzwert nicht erforderlich. Konkret lässt sich unter Rückgriff auf die in die Untersuchung eingebundenen 5.078 GEG-relevanten Nichtwohngebäude mit Angaben zur BGF (vgl. Fußnote 71) mittels Verhältnisschätzung eine durchschnittliche BGF von ca. 1.771 m<sup>2</sup> (Standardfehler: 149 m<sup>2</sup>) ermitteln. Zwar ist es aufgrund der Nichtberücksichtigung der 29 Gebäude mit fehlenden Angaben zur BGF nicht ausgeschlossen, dass dieser Schätzwert systematisch verzerrt ist<sup>72</sup>; allerdings ist die Richtung einer etwaigen Verzerrung unbestimmt.

Multipliziert man nun beide (ungerundeten) Schätzwerte (also 1,981 Mio. mit 1.771 m<sup>2</sup>), kommt man zum gesuchten Ergebnis, hier einer geschätzten BGF aller GEG-relevanten Nichtwohngebäude in Höhe von ca. 3,507 Mrd. m<sup>2</sup>.

Führt man das Beispiel fort und möchte man diesen Schätzwert für eine bestimmte Baualtersklasse ermitteln, kommt als weiterer Verhältnisschätzwert noch der (geschätzte) Baualtersklassenanteil hinzu. Insgesamt ist dann eine Multiplikation der folgenden drei Schätzgrößen notwendig:

- geschätzte Zahl aller GEG-relevanten Nichtwohngebäude
- geschätzter Anteil der GEG-relevanten Nichtwohngebäude in der jeweiligen Baualtersklasse an allen GEG-relevanten Nichtwohngebäuden
- geschätzte mittlere BGF in der jeweiligen Baualtersklasse

Da alle „Eingangsschätzgrößen“ stichprobenbedingt unsicher sind und demzufolge jeweils einen Standardfehler aufweisen, gilt dies auch für das Produkt dieser Schätzwerte, also für die geschätzte BGF der GEG-relevanten Nichtwohngebäude (insgesamt oder einer bestimmten Baualtersklasse).

Um den gesuchten Standardfehler dieses Produkts zu bestimmen (also seinerseits zu schätzen), kam im Rahmen dieser Untersuchung das Gaußsche Fehlerfortpflanzungsgesetz zum Einsatz, das sich in allgemeingültiger Form formelmäßig wie folgt darstellt<sup>73</sup>:

$$(26) \quad \hat{V}(f(\hat{u}_1, \hat{u}_2, \dots, \hat{u}_m)) = \sum_{i=1}^m \left( \frac{\partial f}{\partial u_i}(\hat{u}_i) \right)^2 \hat{V}(\hat{u}_i)$$

$f$  kennzeichnet in Gleichung (26) die jeweils betrachtete Funktion der  $m$  Schätzwerte  $\hat{u}_i$ . Im vorliegenden Fall entsprechen die  $\hat{u}_i$ 's somit der geschätzten

<sup>72</sup> Dies ist dann der Fall, wenn sich die aufgrund fehlender Werte ausgeschlossenen 29 Gebäude in Bezug auf die BGF systematisch von den nicht ausgeschlossenen Untersuchungsgebäuden unterscheiden.

<sup>73</sup> Für eine allgemeine Darstellung und Erläuterung des Gaußschen Fehlerfortpflanzungsgesetzes vgl. z. B. (Hartung 1995 S. 326 ff.) und (Bartsch 2001 S. 562 ff.).

Merkmalssumme ( $\hat{t}_\pi$ ) und (mindestens einem) geschätzten Verhältnis zweier Merkmalssummen ( $\hat{R}$ ), während  $f$  für das Produkt dieser Schätzwerte steht.

Indem man die Wurzel aus (26) zieht, gelangt man zum Schätzwert für den Standardfehler (vgl. auch Unterkapitel 3.3.1).

Für das angeführte Beispiel der BGF-Schätzung beträgt der (geschätzte) Standardfehler für die geschätzte BGF aller GEG-relevanter Nichtwohngebäude somit 0,399 (Mrd. m<sup>2</sup>).<sup>74</sup>

Abschließend sei darauf verwiesen, dass das Gaußsche Fehlerfortpflanzungsgesetz von der Unkorreliertheit der Eingangsgrößen, hier also der  $\hat{u}_i$ , ausgeht und Kovarianzen somit unberücksichtigt bleiben. Aus diesem Grund handelt es sich beim Gaußschen Fehlerfortpflanzungsgesetz um ein Näherungsverfahren für die Schätzung von Standardfehlern.

---

<sup>74</sup> Das entspricht (in Mio. m<sup>2</sup>) der Wurzel aus  $(1.981 \cdot 149)^2 + (1.771 \cdot 152)^2$ , wobei in die tatsächliche Berechnung jeweils ungerundete Größen eingegangen sind.

## 4 Nicht-stichprobenbedingte Fehler

Im Gegensatz zu stichprobenbedingten Fehlern, also der Unsicherheit bei der Übertragung von Stichprobenbefunden auf die Grundgesamtheit, fasst man unter dem Begriff der nicht-stichprobenbedingten Fehler (*nonsampling error*) alle übrigen Fehlerquellen zusammen. Nicht-stichprobenbedingte Fehlerquellen lassen sich (Särndal et al. 1992 S. 16 ff.) zufolge wie folgt unterscheiden, wobei in Klammern die Unterkapitel stehen, in denen nachfolgend diese Fehlerquellen erläutert und ihre Handhabung im vorliegenden Projekt vorgestellt werden:

- Fehler infolge von Nichtbeobachtbarkeit (vgl. Unterkapitel 4.1)
  - Undercoverage (vgl. Unterkapitel 4.1.1)
  - Nonresponse (vgl. Unterkapitel 4.1.2)
- Beobachtungsfehler (vgl. Unterkapitel 4.2)
  - Messfehler (vgl. Unterkapitel 4.2.1)
  - Datenaufbereitungsfehler (vgl. Unterkapitel 4.2.2)

Nicht-stichprobenbedingte Fehler vergrößern genauso wie stichprobenbedingte Fehler die Schätzungenauigkeit. Anders als für die stichprobenbedingten Fehler gibt es allerdings keine umfassende Theorie, um den Einfluss von nicht stichprobenbedingten Fehlern abzuschätzen (vgl. (Lessler, Kalsbeek 1992 S. 5). Dies ist insofern nicht verwunderlich, als nicht-stichprobenbedingte Fehler bisweilen als vermeidbar angesehen werden und sie sich – im Gegensatz zu stichprobenbedingten Fehlern – durch eine „perfekt“ organisierte, keine Kosten und Mühen scheuende Datenerhebung und -aufbereitung prinzipiell vermeiden ließen. In theoretischer Hinsicht mag diese Sichtweise zutreffen, in erhebungspraktischer Hinsicht erscheint sie dagegen naiv. Denn selbst bei einer gründlichen Konzeption und einer straffen Organisation einer Stichprobenbefragung sowie einer sorgfältigen Auswertung der dabei erzielten Befragungsergebnisse lassen sich nicht-stichprobenbedingte Fehler bestenfalls stark eingrenzen, nie aber vollständig unterbinden. Diese Feststellung gilt trotz aller Anstrengungen, nicht-stichprobenbedingte Fehlerquellen zu identifizieren und etwaige Fehler erst gar nicht entstehen zu lassen, auch für die vorliegende Untersuchung. Gleichwohl sollen diese Fehlerquellen in den nachfolgenden Ausführungen benannt und diskutiert werden, zumal ein Teil dieser Fehler – Unit Nonresponse im Screening und in der Breitenerhebung – statistisch modelliert und damit als weiterer Faktor für Schätzungenauigkeiten explizit, d.h. im Standardfehler, berücksichtigt wird.

### 4.1 Fehler infolge von Nichtbeobachtbarkeit

#### 4.1.1 Undercoverage

Wie in Unterkapitel 3.1.4 ausgeführt, können nur solche Nichtwohngebäude in die Untersuchung einbezogen werden, die mit Hausumringen Schnittflächen bilden, die im April 2015 in der Datenbank amtlicher Hausumringe Deutschland enthalten waren. Da sich die Festlegung der Grundgesamtheit an diese „Erreichbarkeitsbedingung“ in der vorliegenden Untersuchung anlehnt, gibt es definitionsgemäß keine Untererfassung und deshalb auch keinen hierdurch hervorgerufenen Fehler.

In dem Maße, in dem der Anspruch erhoben wird, die tatsächlichen Verhältnisse in jüngeren Baualtersklassen oder für das Segment grundflächenmäßig sehr kleiner Nichtwohngebäude und damit in Gebäudesegmenten zu beschreiben, zu

denen die Auswahlgrundlage nur einen lückenhaften Zugang gewährleistet, steigt jedoch die Unsicherheit, da zum quantifizierten Stichprobenbedingten Fehler ein nicht quantifizierbarer Fehler infolge von Undercoverage hinzutritt.

Bei einer inhaltlichen Fokussierung auf den gesamten Nichtwohngebäudebestand oder auf ältere Baualtersklassen dürften die infolge von Untererfassung hervorgerufenen Schätzungenauigkeiten indes kaum ins Gewicht fallen.

#### 4.1.2 Nonresponse

Nonresponse bezeichnet ganz allgemein den vollständigen oder partiellen Ausfall von im Rahmen einer Stichprobenziehung ausgewählten Erhebungs- bzw. Untersuchungseinheiten. Im Fall eines vollständigen Ausfalls spricht man von Unit Nonresponse. Liegen dagegen nur für einzelne Merkmale keine auswertbaren Informationen vor, spricht man von Item Nonresponse.

Nonresponse ist dann vernachlässigbar, wenn er rein zufällig zustande kommt bzw. – anders ausgedrückt – wenn Ausfälle unabhängig von Merkmalsstrukturen sind. In einem solchen Fall spricht man auch von Nonresponse vom Typ MCAR (*missing completely at random*).<sup>75</sup> Die einzige Konsequenz bei dieser Art von Nonresponse ist der damit einhergehende Rückgang des für Auswertungszwecke zur Verfügung stehenden Stichprobenrücklaufs und die damit verbundene Zunahme des Standardfehlers bei Schätzungen. Aber auch wenn Nonresponse von Merkmalsstrukturen abhängt, stellt er noch nicht zwingend ein Problem dar. Denn korreliert der Nonresponse mit beobachtbaren Merkmalsausprägungen und ist er damit durch diese Merkmalsausprägungen erklärbar, lässt er sich bzw. lassen sich die durch ihn hervorgerufenen Verzerrungen korrigieren. In diesem Fall spricht man von Nonresponse vom Typ MAR (*missing at random*).

Da man a priori nicht davon ausgehen kann, dass man es stets mit Nonresponse vom Typ MCAR zu tun hat, besteht die beste Strategie darin, Nonresponse möglichst gering zu halten und dafür zu sorgen, dass bei auftretendem Nonresponse von der Annahme ausgegangen werden kann, er sei vom Typ MAR und damit korrigierbar.

In der vorliegenden Erhebung gab es sowohl Unit als auch Item Nonresponse auf unterschiedlichen Ebenen und in unterschiedlichem Ausmaß. Da mit Unit Nonresponse jedoch anders umgegangen wurde als mit Item Nonresponse, differenzieren die nachfolgenden Ausführungen nach diesen beiden Erscheinungsformen von Nonresponse.

##### Unit Nonresponse

In der vorliegenden Erhebung konnte Unit Nonresponse prinzipiell auf der Ebene der Erhebungsbezirke, im Screening und sowie in den nachfolgenden Erhebungsphasen – der Breiten- sowie der nachgelagerten Tieferhebung – auftreten.

Auf der Ebene der Erhebungsbezirke wurde Unit Nonresponse jedoch dadurch verhindert, dass eine ausreichende Zahl qualifizierter Screener rekrutiert wurde und ausgefallene Screener durch geschultes Ersatzpersonal kompensiert wurden. Alle 500 gezogenen Erhebungsbezirke wurden daher auch tatsächlich begangen.

<sup>75</sup> Für eine formale Definition der verschiedenen Arten von Nonresponse vgl. (Little, Rubin 2002 S. 12).

Unit Nonresponse beschränkte sich daher auf die drei Erhebungsphasen Screening, Breitenerhebung und Tieferhebung, kam aber jeweils in völlig unterschiedlichen Größenordnungen vor. Wie Abbildung 4-1 zu entnehmen ist, konnten ausgehend von der bereinigten Hausumring-Stichprobe<sup>76</sup>, die insgesamt 99.981 Hausumringe beinhaltet, gerade einmal für 58 Stichproben-Hausumringe keine Informationen gewonnen werden, und zwar deshalb, weil diese Hausumringe mangels Begeh- oder wenigstens Einsehbarkeit nicht vor Ort in Augenschein genommen werden konnten und es auch im Nachhinein nicht gelang, anhand von Satellitenaufnahmen Rückschlüsse zu ziehen. Die Ausfallquote ist damit ver-schwindend gering, so dass es als nicht erforderlich erachtet wurde, methodisch ausgereifte und rechenaufwändige Algorithmen zur Aufdeckung und Kompensation etwaiger nonresponsebedingter Verzerrungen einzusetzen. Aufgrund dessen beschränkte sich die vorgenommene Nonresponse-Korrektur darauf, erhebungs-bezirksspezifische Kompensationsfaktoren zu ermitteln und diese allen im Screening identifizierten 89.796 Bauwerken<sup>77</sup> zuzuordnen. Die Kompensationsfaktoren entsprechen dabei dem Quotienten aus dem erhebungsbezirksspezifischen Hausumring-Stichprobenumfang (der im Regelfall bei 200 lag)<sup>78</sup> und dem um die nicht screenbaren Hausumringe dieses Erhebungsbezirks reduzierten Hausumring-Stichprobenumfang.<sup>79</sup> Die Kompensationsfaktoren bewirken im Ergebnis, dass alle Merkmalsausprägungen  $y_k$  und  $z_k$  in den Gleichungen (16), (17), (19), (21), (23) und (24) mit dem jeweiligen erhebungsbezirksspezifischen Kompensationsfaktor multipliziert werden, sofern die hinter diesen Merkmalsausprägungen stehenden Untersuchungsgebäude in Erhebungsbezirken liegen, auf die sich die 58 nicht auf-suchbaren Stichproben-Hausumringe verteilen.

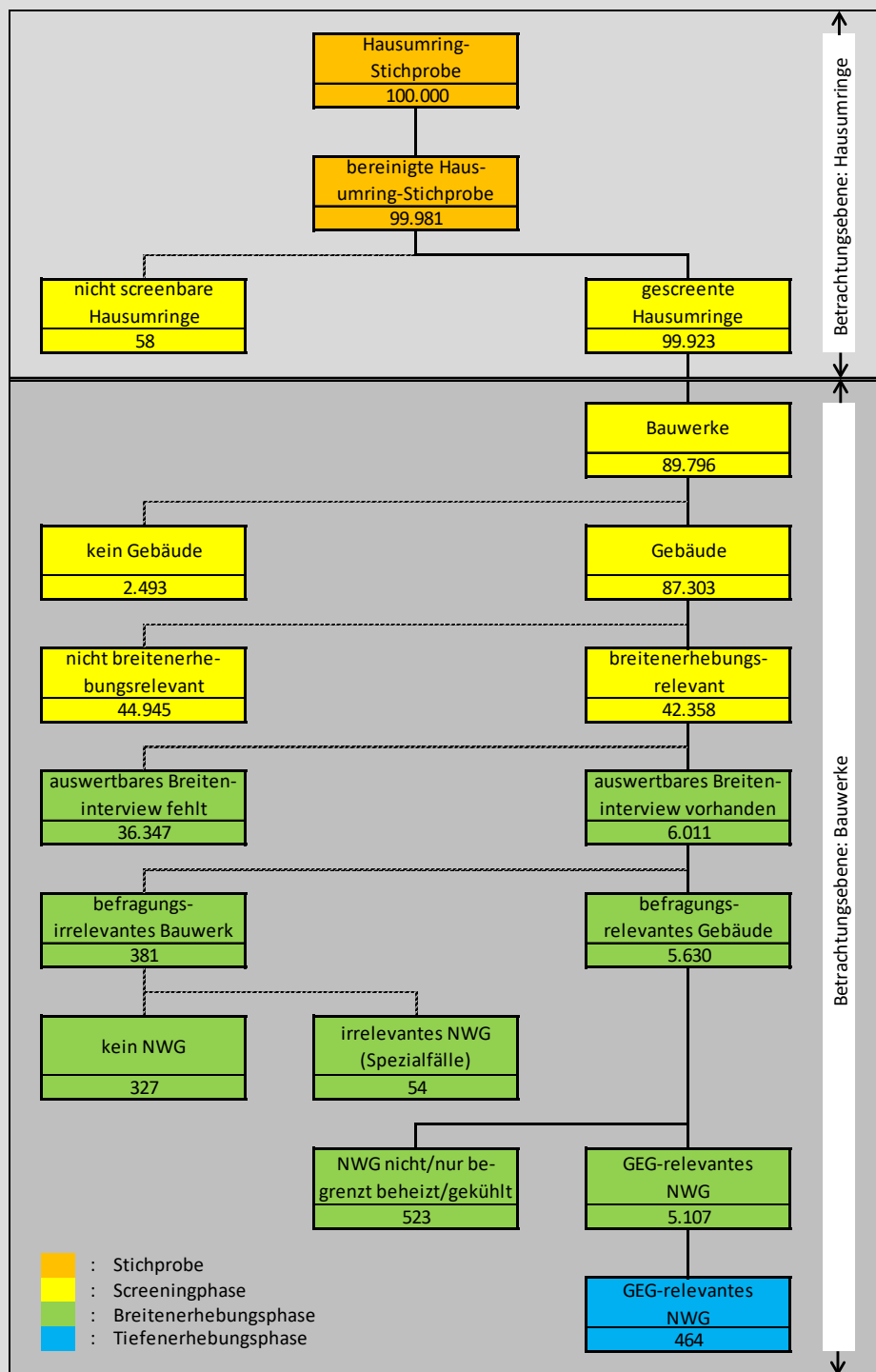
<sup>76</sup> Die Bereinigung der 100.000er Hausumring-Stichprobe war aus technischen Gründen erforderlich und führte zum Ausschluss von insgesamt 19 Hausumringen.

<sup>77</sup> § 2 Abs. 1 Musterbauordnung (MBO) in der aktuellen Fassung von 22.02.2019 zufolge sind bauliche Anlagen (synonym zum Begriff Bauwerk) „mit dem Erdboden verbundene, aus Bauprodukten hergestellte Anlagen; eine Verbindung mit dem Boden besteht auch dann, wenn die Anlage durch eigene Schwere auf dem Boden ruht oder auf ortsfesten Bahnen begrenzt beweglich ist oder wenn die Anlage nach ihrem Verwendungszweck dazu bestimmt ist, überwiegend ortsfest benutzt zu werden“. Und gemäß Abs. 2 sind Gebäude als Unterkategorie von Bauwerken „selbstständig benutzbare, überdeckte bauliche Anlagen, die von Menschen betreten werden können und geeignet oder bestimmt sind, dem Schutz von Menschen, Tieren oder Sachen zu dienen“.

<sup>78</sup> Durch die aus technischen Gründen erforderliche Bereinigung der 100.000er Hausumring-Stichprobe (vgl. auch Fußnote 76) belief sich in einigen wenigen Erhebungsbezirken der Stichprobenumfang auf weniger als 200 Hausumringe.

<sup>79</sup> War beispielsweise in der 200er Hausumring-Stichprobe eines Erhebungsbezirks ein einziger Hausumring nicht screenbar, beläuft sich der Kompensationsfaktor auf  $200/(200-1) \approx 1,005$ .

**Abbildung 4-1: Fallzahlgerüst**



Quelle: Eigene Darstellung.

Abbildung 4-1 zufolge gingen aus den 99.923 gescreenten Hausumringen am Ende 42.358 Gebäude hervor, die in der Screeningphase als breitenerhebungsrelevant



identifiziert wurden, d. h. für die die Durchführung eines Breiteninterviews angestrebt wurde. Bei diesen Gebäuden handelt es sich um Gebäude, die nach dem Erkenntnisstand im Screening mutmaßlich unter die Anwendung des Gebäudeenergiegesetzes (GEG) fallen, also mutmaßlich GEG-relevante Nichtwohngebäude darstellen. Wie bereits am Ende von Unterkapitel 3.1.3 ausgeführt wurde, wurden Zweifelsfälle ebenfalls als breiterhebungsrelevant eingestuft, um der Breiterhebung die finale Entscheidung hinsichtlich der Nichtwohngebäudeeigenschaft wie auch der GEG-Relevanz zu überlassen und um dadurch zu verhindern, dass tatsächlich GEG-relevante Nichtwohngebäude versehentlich von vornherein von einer Berücksichtigung in der Breiterhebung ausgeschlossen werden.

Von diesen 42.358 breiterhebungsrelevanten Gebäuden konnte in 6.011 Fällen ein auswertbares Interview geführt werden. Zu beachten ist dabei, dass ein auswertbares Interview auch dann gegeben ist, wenn es frühzeitig zum Interviewabbruch kam, weil die interviewte Person die Nichtwohngebäudeeigenschaft und/oder die GEG-Relevanz verneinte. Angesichts der Ausführungen im vorangegangenen Absatz, wonach auch Gebäude als breiterhebungsrelevant eingestuft wurden, bei denen Unsicherheit bezüglich der GEG-Relevanz oder sogar bezüglich der Nichtwohngebäudeeigenschaft bestand, verwundert es nicht, dass von den 6.011 Gebäuden mit auswertbarem Breiteninterview immerhin 904 (= 381 + 523) keine GEG-relevanten Nichtwohngebäude sind. In all diesen 904 Fällen wurde auf die Durchführung eines Vollinterviews verzichtet. Stattdessen kam es früher oder später zu einem Interviewabbruch. Ein frühzeitiger Interviewabbruch wurde in Fällen herbeigeführt, bei denen sich herausstellte, dass das betreffende Gebäude entweder kein Nichtwohngebäude ist<sup>80</sup> (327 Fälle) oder aber ein vom GEG ausgeklammerter Spezialfall<sup>81</sup> darstellt (54 Fälle). Stellte sich im Interview heraus, dass ein Nichtwohngebäude überhaupt nicht oder nur in untergeordnetem Umfang unter Einsatz von Energie beheizt oder gekühlt wird, wurde das Interview ebenfalls mangels fehlender GEG-Relevanz<sup>82</sup> abgebrochen, allerdings erst nach dem allgemeinen Frageteil (523 Fälle).

<sup>80</sup> Darunter fallen hauptsächlich Wohngebäude (318 Fälle). In weiteren neun Fällen stellte sich heraus, dass es sich wider Erwarten um ein Bauwerk handelt, bei dem die Gebäudeeigenschaft jedoch fehlt (zu den Definitionen von Bauwerk und Gebäude vgl. Fußnote 77).

<sup>81</sup> Gemeint sind hier die unter § 2 Abs. 2 Nr. 1 bis 2, 4 bis 7 GEG fallenden Gebäude, für die das GEG nur hinsichtlich der Inspektionspflicht für Klimaanlage (§§ 74 bis 78 GEG) gilt. Dabei handelt es sich um „Betriebsgebäude, die überwiegend zur Aufzucht oder zur Haltung von Tieren genutzt werden“ (Nr. 1), „Betriebsgebäude, soweit sie nach ihrem Verwendungszweck großflächig und lang anhaltend offen gehalten werden müssen“ (Nr. 2), „Unterglasanlagen und Kulturräume für Aufzucht, Vermehrung und Verkauf von Pflanzen“ (Nr. 4), „Traglufthallen und Zelte“ (Nr. 5), „Gebäude, die dazu bestimmt sind, wiederholt aufgestellt und zerlegt zu werden, und provisorische Gebäude mit einer geplanten Nutzungsdauer von bis zu zwei Jahren“ (Nr. 6) sowie um „Gebäude, die dem Gottesdienst oder anderen religiösen Zwecken gewidmet sind“ (Nr. 7).

<sup>82</sup> Das GEG gilt für Gebäude, soweit sie unter Einsatz von Energie beheizt oder gekühlt werden (§ 2 Abs. 1 Nr. 1 GEG). Im Umkehrschluss gilt das GEG für unbeheizte und ungekühlte Gebäude somit nicht. Für „sonstige handwerkliche, landwirtschaftliche, gewerbliche, industrielle oder für öffentliche Zwecke genutzte Betriebsgebäude, die nach ihrer Zweckbestimmung auf eine Raum-Solltemperatur von weniger als 12 Grad Celsius beheizt werden oder jährlich weniger als vier Monate beheizt sowie jährlich weniger als zwei Monate gekühlt werden“, gilt das GEG ebenfalls nicht (§ 2 Abs. 2 Nr. 9 GEG). Im Fragebogen wurden diese Fallkonstellationen abgeprüft und für in diesem Zuge identifizierte Nichtwohngebäude kam es nach dem allgemeinen Frageteil zum erwähnten Abbruch des Interviews. Nichtwohngebäude, für die im Interview ein beheizter Flächenanteil  $\leq 10\%$  angegeben wurde, wurden genauso behandelt.

Die zu den Gebäuden mit auswertbarem Breiteninterview komplementäre Menge derjenigen Gebäude, für die ein auswertbares Interview fehlt, ist ungleich größer und beläuft sich auf 36.347 der 42.358 breiterhebungsrelevanten Gebäude. Die Gründe für das Fehlen eines auswertbaren Interviews sind dabei vielfältig. Neben den „klassischen“ Ausfallgründen wie beispielsweise einer Teilnahmeverweigerung aus Zeitgründen kamen weitere hinzu wie etwa die Unmöglichkeit, Kontaktdaten (Telefonnummer, E-Mail-Adresse) zu recherchieren.

Die Zahl von 36.347 breiterhebungsrelevanten Gebäuden mit fehlendem auswertbarem Breiteninterview beziffert somit das Ausmaß von Unit Nonresponse in der Breiterhebung. Um diesen Unit Nonresponse dahingehend zu qualifizieren, ob er systematischer Natur, also nicht vom Typ MCAR ist, und um ihn uno actu zu kompensieren, wurde auf Gebäudemerkmale zurückgegriffen, die im Screening erhoben bzw. generiert wurden. Da – wie in Kürze noch erläutert werden wird – Item Nonresponse im Screening faktisch nicht auftrat, lagen im Regelfall für alle 42.358 breiterhebungsrelevante Gebäude diverse gebäudebeschreibende Informationen vor. Dadurch konnte geprüft werden, ob sich die 36.347 breiterhebungsrelevanten Gebäude mit fehlendem auswertbarem Breiteninterview systematisch von den 6.011 Gebäuden mit auswertbarem Breiteninterview unterscheiden.<sup>83</sup> Konkret gingen in die Prüfung folgende, aus dem Screening hervorgegangenen Gebäudemerkmale (bzw. im Falle einer Nominal- bzw. Ordinalskalierung daraus abgeleitete Dummy-Variablen) ein:

- Primäre Gebäudefunktion (Hauptkategorie) (Screener-Feststellung)
- Baualtersklasse (Screener-Feststellung)
- Modernisierungszustand (Screener-Feststellung)
- Primäre Fassadengestaltung (Screener-Feststellung)
- Vorwiegende Dachform (Screener-Feststellung)
- Mittlere Anzahl der Vollgeschosse (Screener-Feststellung)
- Grundfläche
- Hausumringflächengewichtete Gebäudehöhe
- Gebäudevolumen (= Produkt aus der Grundfläche und der hausumringflächengewichteten Gebäudehöhe)
- Bundesland
- Regionalstatistischer Gemeindetyp
- Anteil der Büronutzung an der Bruttogrundfläche

Die Prüfung erfolgte über eine binär-logistische Regression, bei der die zu erklärende bzw. endogene dichotome Variable darauf abstellte, ob ein auswertbares Breiteninterview vorhanden ist (Ausprägung = 1; 6.011 Fälle) oder nicht (Ausprägung = 0; 36.347 Fälle), während als erklärende bzw. exogene Variablen die aufgelisteten Gebäudemerkmale (bzw. daraus abgeleitete Dummy-Variablen) fungierten. Zur Ermittlung der Schätzgleichung wurde auf die sog. FSTEP-Methode

<sup>83</sup> Einen solchen Vergleich zu ermöglichen, war zwar nicht die zentrale, aber gleichwohl eine wichtige Aufgabe des Screenings (vgl. Unterkapitel 3.1.5).

zurückgegriffen, die das Programmpaket SPSS zur Verfügung stellt<sup>84</sup> und in der am Ende folgende der oben aufgelisteten Merkmale vertreten waren:

- Primäre Gebäudefunktion (Hauptkategorie) (Screener-Feststellung)
- Baualtersklasse (Screener-Feststellung)
- Modernisierungszustand (Screener-Feststellung)
- Primäre Fassadengestaltung (Screener-Feststellung)
- Vorwiegende Dachform (Screener-Feststellung)
- Bundesland
- Regionalstatistischer Gemeindetyp

Der binär-logistische Regressionsansatz bringt neben seiner Eignung zur Prüfung auf statistische Signifikanz der eingebundenen exogenen Variablen den Vorteil mit, dass sich mithilfe der Schätzgleichung auch Wahrscheinlichkeiten, im vorliegenden Fall – hier: Wahrscheinlichkeiten für das Vorliegen eines auswertbaren Breiteninterviews – ableiten lassen. Diese Wahrscheinlichkeiten werden im Folgenden mit  $p_k$  bezeichnet. Im vorliegenden Fall streuen diese Wahrscheinlichkeiten um einen mittleren Wert von 0,161, wobei die Spannweite der Wahrscheinlichkeiten von 0,039 bis 0,435 reicht und die Verteilung leicht rechtsschief ist.

Der Umstand, dass die binär-logistisch ermittelten Wahrscheinlichkeiten  $p_k$  streuen, ist als Beleg dafür zu bewerten, dass der Unit Nonresponse nicht ausschließlich rein zufällig ist (Typ MCAR), sondern dass er in jedem Fall (auch) mit beobachtbaren Merkmalsausprägungen korreliert (Typ MAR). Wie das Pseudo-Bestimmtheitsmaß Nagelkerkes Pseudo- $R^2$ , das sich im vorliegenden Fall auf 0,038 und damit auf einen Wert deutlich unter dem Maximalwert von 1 beläuft, belegt, kann das geschätzte Modell die (beiden) Ausprägungen der endogenen Variable erwartungsgemäß nicht vollständig erklären. Offenkundig gibt es somit noch andere, nicht beobachtete Einflussfaktoren für die Teilnahmebereitschaft bzw. – genauer – dafür, dass zu einem breiterhebungsrelevanten Gebäude ein auswertbares Breiteninterview generiert werden konnte. Es wird unterstellt, dass diese Einflussfaktoren unsystematischer Natur in dem Sinne sind, dass sie nicht mit Gebäudemerkmalen korrelieren, die Gegenstand späterer Ergebnisauswertungen sind.

Indem die regressionsanalytisch abgeleiteten Wahrscheinlichkeiten  $p_k$  bei der Schätzung von Populationsparametern adäquat berücksichtigt werden, gelingt es, verzerrendem Unit Nonresponse wirkungsvoll gegenzusteuern. Die Berücksichtigung der Wahrscheinlichkeiten  $p_k$  erfolgt dabei durch eine entsprechende Modifizierung der Formeln zur Schätzung von Merkmalssummen, von Verhältnissen zweier Merkmalssummen sowie von Standardfehlern.

Was die Merkmalssumme  $t$  angeht, lässt sich diese bei Berücksichtigung von Unit Nonresponse durch eine Modifizierung des HT-Schätzers  $\hat{t}_\pi$  aus Gleichung (16) wie folgt schätzen (vgl. (Stenger 1986 S. 209)):

<sup>84</sup> Die FSTEP-Methode (FSTEP steht für „Forward stepwise“) startet mit einem Modell ohne exogene Variablen und nimmt in jedem Schritt dasjenige Merkmal in die Modellgleichung auf, welches bei Aufnahme mit dem geringsten p-Wert (sofern kleiner als 0,05) einhergeht. Nach jedem Aufnahmeschritt werden die bereits vorhandenen Merkmale auf Ausschluss geprüft, wobei das Merkmal mit dem höchsten p-Wert (sofern größer als 0,06) aus dem Modell ausgeschlossen wird. Die endgültige Modellgleichung ist erreicht, wenn keine exogene Variable mehr aufgenommen werden kann, die das Aufnahmekriterium (p-Wert < 0,05) erfüllt.

$$(27) \quad \hat{t}_\pi = \sum_{\tilde{U}} L_k L_k^* \frac{y_k}{\pi_k p_k} \quad \forall \pi_k > 0, p_k > 0 \text{ und } k \in \tilde{U} \subset U \text{ mit}$$

$L_k^*$  : Indikatorvariable, die darauf abstellt, ob von der Untersuchungseinheit  $k$  (d. h. von Nichtwohngebäude  $k$ ) ein auswertbares Breiteninterview vorliegt

Dass Gleichung (27) anders als Gleichung (16) auf die Teilmenge  $\tilde{U}$  der Grundgesamtheit fokussiert ( $\tilde{U} \subset U$ ), ist dem Umstand geschuldet, dass nur ein Teil der gescreenten Gebäude breitenenerhebungsrelevant ist und folglich nur für diese Teilpopulation Populationsparameter in Bezug auf in der Breitenenerhebung erhobene Merkmale geschätzt werden können.<sup>85</sup> Je nach Auswertungsinteresse muss  $\tilde{U}$  noch weiter eingegrenzt werden, unter anderem dann, wenn nur GEG-relevante Nichtwohngebäude oder wiederum Teilmengen davon (z. B. GEG-relevante Nichtwohngebäude bestimmter Baualtersklassen) betrachtet werden.

Da die Indikatorvariable  $L_k^*$  nur für die hinter den 6.011 auswertbaren Breiteninterviews stehenden Bauwerke bzw. Nichtwohngebäude den Wert 1 und andernfalls den Wert 0 annimmt, lässt sich der HT-Schätzer  $\hat{t}_\pi$  aus Gleichung (27) auch in anderer Form darstellen:

$$(28) \quad \hat{t}_\pi = \sum_{\bar{s}} \frac{y_k}{\pi_k p_k} \quad \forall \pi_k > 0, p_k > 0 \text{ und } k \in \bar{s} \subset \tilde{s} \subset s \text{ mit}$$

$\bar{s}$  : Menge der Untersuchungseinheiten (d. h. also der Nichtwohngebäude), für die ein auswertbares Breiteninterview vorliegt,

wobei  $\tilde{s}$  hier auf die 42.358 im Screening als breitenenerhebungsrelevant eingestuftten Gebäude abstellt. Gleichung (28) ist somit das Gegenstück zu Gleichung (17) aus Unterkapitel 3.3.1, wo von der Existenz von Unit Nonresponse noch abgesehen wurde.

Durch die Berücksichtigung von Unit Nonresponse und die dadurch bewirkte Modifizierung des HT-Schätzers  $\hat{t}_\pi$  und unter der Annahme, dass der Unit Nonresponse vom Typ MCAR ist, ändert sich nichts an der Erwartungstreue bzw. Unverzerrtheit dieses Schätzers (vgl. ähnlich ebenda).

Nimmt man an, dass das Vorliegen auswertbarer Breiteninterviews für zwei beliebige breitenenerhebungsrelevante Nichtwohngebäude  $k$  und  $l$  unabhängige Ereignisse sind, dass also  $p_{kl} = p_k p_l$  ist<sup>86</sup>, gilt für die Varianz des HT-Schätzers  $\hat{t}_\pi$  aus Gleichung (27) bzw. (28)

$$(29) \quad V(\hat{t}_\pi) = \sum \sum_{\tilde{U}} (\pi_{kl} - \pi_k \pi_l) \frac{y_k y_l}{\pi_k \pi_l} + \sum_{\tilde{U}} \pi_k \frac{1 - p_k}{p_k} \left( \frac{y_k}{\pi_k} \right)^2 \quad \forall \pi_k, \pi_l, p_k > 0 \text{ und } k, l \in \tilde{U},$$

<sup>85</sup> Bei den im Screening als breitenenerhebungsrelevant eingestuften Nichtwohngebäuden handelt es sich daher um eine sog. *domain* (vgl. letzter Absatz von Unterkapitel 3.3.1).

<sup>86</sup> Die Annahme läuft letztlich darauf hinaus zu unterstellen, dass die Entscheidung für die Teilnahme am Breiteninterview für jedes breitenenerhebungsrelevante Nichtwohngebäude losgelöst von der Teilnahmeentscheidung für ein anderes breitenenerhebungsrelevantes Nichtwohngebäude erfolgt. Sofern es sich um breitenenerhebungsrelevante Nichtwohngebäude handelt, deren Ansprechpartner nicht auch Ansprechpartner für andere breitenenerhebungsrelevante Nichtwohngebäude sind, ist diese Annahme unmittelbar einsichtig. In Fällen von Personen, die für mehrere breitenenerhebungsrelevante Nichtwohngebäude als Ansprechpartner fungieren, impliziert die Annahme, dass der jeweilige Ansprechpartner seine Teilnahmeentscheidung für jedes „seiner“ breitenenerhebungsrelevanten Nichtwohngebäude isoliert trifft.

wobei der zweite Summand die Varianzvergrößerung infolge von Unit Non-response angibt (vgl. ebenda, S. 210).

Ein erwartungstreuer Schätzer für  $V(\hat{t}_\pi)$ , der auch in der vorliegenden Untersuchung zum Einsatz kam, ist gegeben durch

$$(30) \quad \hat{V}(\hat{t}_\pi) = \sum \sum_{\bar{s}} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k p_k} \frac{y_l}{\pi_l p_l} + \sum_{\bar{s}} \pi_k \frac{1 - p_k}{p_k^2} \left( \frac{y_k}{\pi_k} \right)^2$$

$\forall \pi_k, \pi_l, \pi_{kl}, p_k, p_l > 0 \text{ und } k, l \in \bar{s}$

(ähnlich vgl. ebenda).

Für die Schätzung des Verhältnisses zweier Merkmalssummen, die – im Fall von Abwesenheit von Unit Nonresponse – in Unterkapitel 3.3.2 erläutert wurde, gilt bei Berücksichtigung von Unit Nonresponse im Wesentlichen dasselbe. So verändert sich der Verhältnisschätzer aus Gleichung (21) zu

$$(31) \quad \hat{R} = \frac{\hat{t}_{y\pi}}{\hat{t}_{z\pi}} = \frac{\sum_{\bar{s}} y_k / \pi_k p_k}{\sum_{\bar{s}} z_k / \pi_k p_k},$$

während die einzige Änderung von Gleichung (22) – der Formel für die approximative Varianz  $AV(\hat{R})$  – darin besteht, dass  $U$  durch  $\tilde{U}$  zu ersetzen ist.

Die Formel für den Varianzschätzer stellt sich genauso wie in Gleichung (23) dar; allerdings sind jetzt  $\hat{V}(\hat{t}_{y\pi})$  bzw.  $\hat{V}(\hat{t}_{z\pi})$  wie in Gleichung (30) und  $\hat{R}^2$  unter Rückgriff auf Gleichung (31) zu berechnen, während für den Kovarianzschätzer aus Gleichung (24) jetzt

$$(32) \quad \hat{C}(\hat{t}_{y\pi}, \hat{t}_{z\pi}) = \sum \sum_{\tilde{U}} L_{kl} L_{kl}^* \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k p_k} \frac{z_l}{\pi_l p_l}$$

$\forall \pi_k, \pi_l, \pi_{kl}, p_k, p_l > 0 \text{ und } k, l \in \tilde{U}$

mit

$L_{kl}^*$  : Indikatorvariable, die darauf abstellt, ob von den beiden Untersuchungseinheiten  $k$  und  $l$  (d. h. von den beiden Nichtwohngebäuden  $k$  und  $l$ ) auswertbare Breiteninterviews vorliegen

gilt.

Alternativ lässt sich Gleichung (32) auch ausdrücken als

$$(33) \quad \hat{C}(\hat{t}_{y\pi}, \hat{t}_{z\pi}) = \sum \sum_{\bar{s}} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k p_k} \frac{z_l}{\pi_l p_l}$$

$\forall \pi_k, \pi_l, \pi_{kl}, p_k, p_l > 0 \text{ und } k, l \in \bar{s}$

Die Erwartungstreue dieses Schätzers kann unter Zugrundelegung der bereits erwähnten Unabhängigkeitsannahme wie folgt bewiesen werden:

$$\begin{aligned} \text{Beweis: } E\hat{C}(\hat{t}_{y\pi}, \hat{t}_{z\pi}) &= E_1 E_2 \sum \sum_{\tilde{U}} L_{kl} L_{kl}^* \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k p_k} \frac{z_l}{\pi_l p_l} \\ &= E_1 \sum \sum_{\tilde{U}} L_{kl} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k p_k} \frac{z_l}{\pi_l p_l} E_2 L_{kl}^* \\ &= E_1 \sum \sum_{\tilde{U}} L_{kl} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k p_k} \frac{z_l}{\pi_l p_l} p_{kl} \\ &= E_1 \sum \sum_{\tilde{U}} L_{kl} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k p_k} \frac{z_l}{\pi_l p_l} p_k p_l \\ &= \sum \sum_{\tilde{U}} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k} \frac{z_l}{\pi_l} E_1 L_{kl} \end{aligned}$$

$$\begin{aligned}
 &= \sum \sum_{\tilde{U}} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k z_l}{\pi_k \pi_l} \pi_{kl} \\
 &= \sum \sum_{\tilde{U}} (\pi_{kl} - \pi_k \pi_l) \frac{y_k z_l}{\pi_k \pi_l} \\
 &= C(\hat{t}_{y\pi}, \hat{t}_{z\pi})
 \end{aligned}$$

□

Unit Nonresponse gab es auch beim Übergang von der Breiten- zur Tieferhebung. Konkret konnten „nur“ für 464 der 5.107 GEG-relevanten Nichtwohngebäude mit auswertbaren Vollinterviews ebenfalls auswertbare Tiefeninterviews realisiert werden (vgl. Abbildung 4-1). Das auf den ersten Blick beachtliche Ausmaß von Unit Nonresponse lag dabei angesichts des Aufwandes, der den Interviewpartnern im Zuge der Durchführung eines Tiefeninterviews entstand und ihnen vorab auch kommuniziert wurde, im Bereich des Erwarteten. Anders als beim Unit Nonresponse, der in der Breitenhebung aufgetreten ist, wurde allerdings darauf verzichtet, den Unit Nonresponse der Tieferhebung statistisch zu modellieren, d.h. etwaige Verzerrungen mittels eines binär-logistischen Regressionsansatzes aufzudecken und entsprechend zu kompensieren. Der Grund für diese Ungleichbehandlung lag darin, dass angesichts der großen Heterogenität im Nichtwohngebäudesektor ein „Repräsentativitätsanspruch“ von nur wenigen Hundert Untersuchungsgebäuden, die in die Tieferhebung eingebunden waren, verneint wurde. Die Tiefeninterviews sind dadurch gewiss nicht wertlos, sondern eignen sich für vielfältige Untersuchungen wie beispielsweise der Analyse zur Diskrepanz von gemessenem Energieverbrauch und berechnetem Energiebedarf. Den Anspruch, die Grundgesamtheit aller GEG-relevanten Nichtwohngebäude mit vertretbarer statistischer Unsicherheit beschreiben zu wollen, können die wenigen hundert „Tieferhebungsgebäude“ indes nicht erfüllen.

### Item Nonresponse

Item Nonresponse gab es in der vorliegenden Erhebung nur in verhältnismäßig geringem Umfang. Im Screening konnte Item Nonresponse dank der sorgfältigen Rekrutierung, Schulung und Betreuung des mit dem Screening betrauten Personals, der eingesetzten Erhebungs-App sowie dank der Nachkontrolle der vor Ort erhobenen Daten durch das Kontrollteam des Verbundpartners BUW-ÖPB sogar vollständig verhindert werden, d. h. ausnahmslos für alle 87.303 im Screening identifizierten Gebäude liegen die zu erhebenden Merkmalsausprägungen vor.

Was die Breitenhebung angeht, war der dort eingesetzte Fragebogen durch eine ganze Reihe von Maßnahmen so gestaltet, dass fehlende Angaben der Befragten und damit Item Nonresponse auf ein Mindestmaß zurückgedrängt werden konnte. Von entscheidender Bedeutung war dabei, auf halboffene und erst recht auf offene Fragen weitgehend zu verzichten sowie bei den geschlossenen Fragen die Antwortoptionen „weiß nicht“ oder „keine Angabe“ – von Ausnahmen abgesehen – erst gar nicht vorzusehen.<sup>87</sup> Und in den Fällen mit eingeräumter

<sup>87</sup> Zu erwähnen ist in diesem Zusammenhang vor allem die dreistufige Abfrage in Bezug auf viele metrische Merkmale wie beispielsweise das Baujahr oder den beheizten und/oder gekühlten Flächenanteil des Gebäudes: Zunächst wurden die Befragten gebeten, eine genaue (metrische) Angabe zu machen, sie konnten aber auch „weiß nicht“ antworten. Im letzten Fall wurden sie zur Angabe eines (ebenfalls metrischen) Schätzwerts aufgefordert, wobei wiederum die Angabe „weiß nicht“ möglich war. Machten die Befragten erneut von dieser Option Gebrauch, bekamen sie Wertklassen ohne „Weiß nicht“-Option vorgelegt. Durch diese Form der Abfrage wurde gewährleistet, dass für grundsätzlich jedes Gebäude eine Wertangabe vorhanden ist, wenn auch nur in klassifizierter Form. Im Zuge der Datenaufbereitung wurden die Angaben ungeachtet der Stufe, auf der sie gemacht



Möglichkeit, eine Angabe zu unterlassen, beugte das Grundprinzip, einfache Fragen zu stellen, Fachbegriffe allgemeinverständlich zu erläutern und auf möglicherweise als sensibel empfundene Fragen (z. B. dazu, ob auf dem Grundstück ein Grundpfandrecht in Gestalt einer Hypothek oder einer Grundschuld lastet) zu verzichten, Item Nonresponse wirkungsvoll vor. Hinzu kam eine layouttechnisch ansprechende, übersichtliche und mit zahlreichen Piktogrammen versehene Fragebogengestaltung, denn in einem Drittel der Fälle erfolgte die Befragung nicht interviewergestützt auf telefonischem Wege (sog. CATI (*Computer Assisted Telephone Interview*)), sondern online ohne Beisein eines Interviewers. Und auch die unternommenen Anstrengungen, um einen kompetenten und auskunftsfreudigen Ansprechpartner zu gewinnen, dienten nicht zuletzt dazu, Item Nonresponse in tolerierbaren Grenzen zu halten.

Auch in Bezug auf die Tieferhebung wurde Item-Nonresponse weitgehend vermieden. Dazu trug nicht zuletzt bei, dass die Daten vor Ort von kompetenten Energieberatern erhoben wurden, die fast ausnahmslos gemäß der Energieeffizienz-Experten-Liste für Nichtwohngebäude der Deutschen Energieagentur (dena) zertifiziert waren und die darüber hinaus zur Erfüllung der an sie gestellten Anforderungen bei der Datenaufnahme gesondert geschult wurden. Auch der Umstand, dass zur Datenaufnahme ein Tablet-PC mit einheitlicher Konfiguration und eine Erhebungsmaske auf Basis von Microsoft-EXCEL zum Einsatz kamen, trug wirkungsvoll zur Eindämmung fehlender Angaben bei. Dass in Einzelfällen dennoch Item Nonresponse zu verzeichnen war, war z. B. darauf zurückzuführen, dass Technikräume nicht zugänglich waren, weil der Ansprechpartner vor Ort nicht den passenden Schlüssel zur Verfügung hatte. In diesen Fällen gibt es aber eine klare Kennzeichnung und einen nachvollziehbaren Grund.

Item Nonresponse in der Breiten- und Tieferhebung entstand nicht nur durch unterbliebene Angaben der Befragten, sondern auch aus zwei anderen Gründen.

Zum einen erzeugte die Aufbereitung der erhobenen Rohdaten in wenigen Fällen Item Nonresponse, nämlich dadurch, dass Befragtenangaben zu bestimmten Sachverhalten gelöscht wurden, weil sie als sicher falsch identifiziert werden konnten, ohne dass es aber gleichzeitig gelang, die offenkundig falsche Angabe durch eine richtige zu ersetzen. Ein Beispiel hierfür sind Fälle, in denen Befragte auf die Frage: „Welche Dämmstoffdicke in cm liegt überwiegend vor?“ mit „280“ oder „150“ geantwortet haben. Solche Dämmstärken in cm sind mit an Sicherheit grenzender Wahrscheinlichkeit aus technischen Gründen auszuschließen. Entweder unterlagen Befragte einem Missverständnis und meinten, Dämmstärken in mm angeben zu sollen, oder die Antwort beruhte auf einem Tippfehler bei der Beantwortung des Online-Fragebogens. Angaben dieser Art, bei denen die richtige Antwort unbekannt war, wurden gelöscht und der dadurch entstehende fehlende Wert wurde durch eine entsprechende Missing-Codierung gekennzeichnet.

Zum anderen wurden bei acht der 6.011 breiterhebungsrelevanten Gebäude fast alle in der Breiterhebung generierten Merkmale gelöscht. Diese Löschung

---

wurden, in einer gemeinsamen (metrischen oder klassifizierten, d. h. ordinal skalierten) Variable zusammengefasst, die nahezu frei von Antwortausfällen war. Die verbleibenden wenigen Fälle mit Antwortausfällen erklären sich mit nachträglichen Plausibilisierungs- und Editierungsmaßnahmen, in deren Zuge Einträge als unplausibel verworfen wurden und die dahinterstehenden Gebäude die Merkmalsausprägung „unbekannt“ erhielten.



war dem Umstand geschuldet, dass die betreffenden Gebäude durch „unglückliche“ Umstände „kombinierte“ Hochrechnungsfaktoren<sup>88</sup> von jeweils über einem Prozent des gesamten Hochrechnungsvolumens aufwiesen und dadurch Auswertungen dominiert hätten. Um dies zu unterbinden, erfolgte die Merkmalslöschung. Davon ausgenommen war lediglich das Merkmal zur Zuordnung eines breiterehebungsrelevanten Gebäudes zu den verschiedenen grundlegenden Gebäudetypen, um hierfür möglichst verzerrungsfreie Merkmalssummen schätzen zu können.<sup>89</sup>

Item Nonresponse gab es darüber hinaus auch bei Merkmalen, die nicht auf dem Befragungswege erhoben wurden, sondern die auf sekundärstatistischem Wege ermittelt und den Untersuchungsgebäuden zugespielt wurden. Konkret handelt es sich hierbei um Merkmale zur Beschreibung der Gebäudehüllfläche.

In die spätere Auswertung von Merkmalen mit Item Nonresponse wurden aufgrund des regelmäßig geringen Umfangs der fehlenden Angaben stets nur Gebäude mit Angaben eingebunden. Der damit einhergehende Verzicht auf entsprechende Korrekturalgorithmen impliziert die Annahme, dass der vorhandene Item Nonresponse nicht verzerrend in dem Sinne ist, dass die dahinterstehenden Gebäude in Bezug auf die Auswertungsmerkmale andere Strukturen aufweisen als die Gebäude mit Angaben. Aufgrund des Umstandes, dass – von wenigen Ausnahmen abgesehen – fast alle erhobenen bzw. daraus abgeleiteten oder anderweitig zugespielten Merkmale von Item Nonresponse betroffen waren und dadurch fehlende Werte aufwiesen, wurde bei fast allen Merkmalssummenschätzungen wie in Unterkapitel 3.3.3 beschrieben vorgegangen.

## 4.2 Beobachtungsfehler

Bei sog. Beobachtungsfehlern (*error in observations*) entspricht die letztlich aufgezeichnete und für die Schätzung herangezogene Merkmalsausprägung nicht der Wahrheit (Särndal et al. 1992 S. 16). Beobachtungsfehler lassen sich in Messfehler (*measurement error*) und Datenaufbereitungsfehler (*processing error*) unterscheiden. Beide Unterkategorien von Beobachtungsfehlern wurden im Rahmen dieser Untersuchung nicht mathematisch modelliert, weil sie nicht quantifiziert werden konnten<sup>90</sup>. Stattdessen wurden die gesamte Erhebungsorganisation sowie die Erhebungsinstrumente darauf ausgerichtet, die Aufzeichnung und Speicherung wahrheitswidriger Merkmalsausprägung auf ein Minimum zu reduzieren.

<sup>88</sup> Der kombinierte Hochrechnungsfaktor entspricht dem Kehrwert des Produkts aus der Ziehungswahrscheinlichkeit  $\pi_k$  und der (binär-logistisch ermittelten) Teilnahmewahrscheinlichkeit  $p_k$  und gibt an, für wie viele Gebäude der Grundgesamtheit das Stichprobengebäude stellvertretend steht. Da es sich bei beiden Einzelwahrscheinlichkeiten um gebäudeindividuelle Werte handelt, war nicht ausgeschlossen, dass in Einzelfällen sehr geringe Werte sowohl für  $\pi_k$  als auch für  $p_k$  zusammentrafen.

<sup>89</sup> Diese Merkmalssummen sind in Tabelle 3-6 ausgewiesen. Vgl. auch die diesbezüglichen Ausführungen in Unterkapitel 3.3.3.

<sup>90</sup> Eine Quantifizierung hätte mangels einschlägigen Erfahrungswissens experimenteller Verfahren bedurft, bei der für eine Teilmenge von Fällen die von den Screenern bzw. Befragten gemachten und im Zuge der Datensatzerstellung aufbereiteten Angaben auf ihre Richtigkeit hin überprüft werden. Der damit einhergehende Aufwand wurde jedoch in keinem angemessenen Verhältnis zum daraus entstehenden Ertrag gesehen.

#### 4.2.1 Messfehler

Messfehler entstehen typischerweise durch Falschangaben von Befragten oder durch Falschaufzeichnungen des Erhebungspersonals, im vorliegenden Fall also des Screeningpersonals, der Interviewer in der Breitenerhebung (sofern das Breiteninterview telefonisch geführt wurde) und der Energieberater in der Tiefenerhebung. Darüber hinaus können sich im vorliegenden Projekt auch bei Merkmalen, die aus sekundärstatistischem Datenmaterial abgeleitet und den Untersuchungsgebäuden nachträglich zugespielt wurden (z.B. Merkmale, die die Gebäuhüllfläche beschreiben), Messfehler eingeschlichen haben.

Durch vielfältige Anstrengungen, beginnend mit dem Einsatz qualifizierten Erhebungspersonals über die sorgfältige Gestaltung der Erhebungsunterlagen bis hin zur Suche von kompetenten Gebäudeansprechpartnern, wurde versucht, Messfehler weitestgehend zurückzudrängen. Und auch die Datenaufbereitung im Nachgang jeder Erhebungsphase war darauf ausgerichtet, Messfehler aufzudecken und möglichst zu korrigieren. Wie bei jeder anderen Erhebung ist allerdings auch bei der vorliegenden Untersuchung nicht davon auszugehen, dass die erhobenen Daten völlig frei von Messfehlern sind bzw. dass es im Zuge der nachgelagerten Datenaufbereitung gelang, alle Messfehler aufzudecken und zu korrigieren.

#### 4.2.2 Datenaufbereitungsfehler

Fehler können auch im Zuge der Datenaufbereitung entstehen, nämlich indem mutmaßlich falsche Angaben durch Werte überschrieben werden, die ihrerseits falsch sind. Um nicht Gefahr zu laufen, seltene, aber dennoch wahrheitsgetreue Gebäudemerkmale im Zuge der Datenaufbereitung irrtümlicherweise als Falschangaben zu identifizieren und durch vermutete richtige, in Wirklichkeit aber falsche Ausprägungen zu ersetzen, beschränkten sich solche Überschreibungen von Merkmalsausprägungen daher auf wohlbegründete Ausnahmefälle.

Wurde zu einem eingeschossigen Gebäude beispielsweise eine auffallend große Nettoraumfläche (NRF) von 600.000 m<sup>2</sup> angegeben, ergab die Nachprüfung anhand von Gebäudefotos, Schrägluftbildern und Nachmessungen mit ArcGIS-Online aber, dass die NRF ca. 60.000 m<sup>2</sup> betragen dürfte, wurde die offensichtlich um den Faktor 10 zu hoch eingetragene NRF durch diesen Wert ersetzt. Diese Überschreibung der Befragtenangabe selbst ist natürlich auch mit dem Risiko von Fehlern behaftet, allein schon deshalb, weil es sich bei der korrigierten Merkmalsausprägung (60.000 m<sup>2</sup>) um einen gerundeten Wert handelt, der sehr wahrscheinlich nicht der exakten NRF entspricht. Eine weitere potentielle Quelle von Datenaufbereitungsfehlern entstand durch die Auswertung der Angaben von Befragten in Freitextfeldern, die der in der Breitenerhebung verwendete Fragebogen an mehreren Stellen vorsah. Die textlichen Eintragungen der Befragten waren nicht selten unpräzise und deshalb interpretationsbedürftig. Folglich ist nicht auszuschließen, dass trotz aller Sorgfalt solche Eintragungen falsch interpretiert wurden und dadurch falsche Merkmalsausprägungen vergeben wurden.

Aufgrund der behutsamen Vorgehensweise bei der Überschreibung erhobener Daten dürften Datenaufbereitungsfehler im vorliegenden Projekt jedoch kaum ins Gewicht fallen, insbesondere auch deshalb, weil auf den Einsatz von Imputationsverfahren zur Beseitigung fehlender Werte fast ausnahmslos verzichtet wurde. Denn beim Imputieren besteht (fast) immer die Gefahr, Fehler infolge von Item

Nonresponse durch Datenaufbereitungsfehler zu ersetzen (vgl. (Lessler, Kalsbeek 1992 S. 12).

## 5 Literatur

- Bartsch, Hans-Jochen (2001). *Taschenbuch mathematischer Formeln*, 19. Aufl.
- Buttler, Günter; Fickel, Norman (2002). *Statistik mit Stichproben*.
- Cischinsky, Holger; Diefenbach, Nikolaus (2018). Datenerhebung zu den energetischen Merkmalen und Modernisierungsraten im deutschen und hessischen Wohngebäudebestand. *IWU, Darmstadt 2018*.
- Cischinsky, Holger; Kirchner, Joachim; Vaché, Martin; Rodenfels, Markus; Nuss, Galina (2015). *Privateigentümer von Mietwohnungen in Mehrfamilienhäusern*. BBSR-Online-Publikation, Nr. 02/2015.
- Hartmann, André; Behnisch, Martin; Hecht, Robert; Meinel, Gotthard; Schorcht, Martin; Schwarz, Steffen (2020). *Teilbericht Gebäudemerkmale*. (ENOB:dataNWG Teilbericht Nummer E 1.4.3). Dresden: IÖR.
- Hartung, Joachim (1995). *Statistik*, 10. Aufl.
- Lessler, Judith T.; Kalsbeek, William D. (1992). *Nonsampling Error in Surveys*.
- Little, Roderick J. A.; Rubin, Donald B. (2002). *Statistical Analysis with Missing Data*, 2nd edition.
- Särndal, Carl-Erik; Swensson, Bengt; Wretman, Jan (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.
- Schorcht, Martin; Hartmann, André; Behnisch, Martin; Hecht, Robert; Meinel, Gotthard; Schwarz, Steffen (2021). *Teilbericht Erhebungsbezirke*. (ENOB:dataNWG Teilbericht Nummer E 1.4.4). Dresden: IÖR.
- Stenger, Horst (1986). *Stichproben*. Heidelberg, Wien: Physica-Verlag.