

REGIONALE AUSWERTUNG VON UNTERNEHMENSSTATISTIKEN: METHODEN UND ANWENDUNGEN IM KONTEXT DER SMALL-AREA-STATISTIK

Julia Manecke

➤ **Schlüsselwörter:** Kleinräumige Schätzung – Unternehmensstatistiken – Schichtwechsler – Neugewichtung – Frame Errors

ZUSAMMENFASSUNG

Regional und inhaltlich tief gegliederte Ergebnisse von Unternehmensstatistiken sind zunehmend von großem Interesse. Eine entsprechende Disaggregation von Unternehmensstatistiken und die entsprechenden kleinräumigen Schätzungen mithilfe von sogenannten Small-Area-Verfahren unterliegen jedoch verschiedensten Herausforderungen. Dabei sind die in der Realität häufig auftretenden Inkonsistenzen zwischen Erhebungsergebnissen und Unternehmensregistern von besonderer Relevanz und entsprechend zu berücksichtigen. Anhand einer Monte-Carlo-Simulationsstudie wird in dieser Arbeit die Fragestellung untersucht, inwieweit verschiedene Small-Area-Verfahren sowie entwickelte Neugewichtungsmethoden die Schätzgüte bei entsprechenden Inkonsistenzen verbessern können.

➤ **Keywords:** *small area estimation – business statistics – stratum jumpers – reweighting – frame errors*

ABSTRACT

The interest in detailed results of business statistics, broken down by topic and region, has increased significantly. However, an adequate disaggregation of business statistics and the respective estimations by means of small area estimation approaches pose various challenges. In this context, the inconsistencies between survey results and business registers, which are often encountered in reality, are of particular relevance and need to be considered appropriately. On the basis of a Monte Carlo simulation, this thesis examines the extent to which different small area methods and newly developed reweighting approaches can achieve improvements in the estimation quality when dealing with the associated inconsistencies.



Julia Manecke

hat Survey Statistics (M.Sc.) an der Universität Trier studiert und auch das Zertifikat European Master in Official Statistics (EMOS) erlangt. Im Rahmen des EMOS-Programms absolvierte sie einen Forschungsaufenthalt im Hessischen Statistischen Landesamt. Aus dieser Kooperation ist ihre Masterarbeit zum Thema „Regionale Auswertung von Unternehmensstatistiken: Methoden und Anwendungen im Kontext der Small Area-Statistik“ entstanden, für die sie mit dem Gerhard-Fürst-Preis 2018 in der Kategorie „Master-/Bachelorarbeiten“ ausgezeichnet wurde. Seit Mai 2017 ist sie Doktorandin an der Professur für Wirtschafts- und Sozialstatistik an der Universität Trier.

1

Einleitung

In den letzten Jahren ist die Nachfrage nach regional und inhaltlich tief gegliederten Ergebnissen von Unternehmensstatistiken deutlich gestiegen. Das große Nutzerinteresse ist unter anderem darauf zurückzuführen, dass belastbare Daten bei ökonomischen Entscheidungen oder der Planung von regional- und strukturpolitischen Maßnahmen benötigt werden. Oftmals ist das Stichprobendesign einer Erhebung jedoch aufgrund maximal zulässiger Stichprobenumfänge lediglich auf eine zuverlässige designbasierte Schätzung auf Landes- oder Bundesebene ausgelegt. Wenn zusätzliche Schätzwerte für räumlich oder inhaltlich untergliederte Teilgesamtheiten ermittelt werden sollen, können unzureichende Stichprobengrößen zu inakzeptabel hohen Schätzvarianzen führen. Eine Teilgesamtheit, in welcher der Stichprobenumfang nicht für eine direkte designbasierte Schätzung von hinreichender Präzision ausreicht, wird als Small Area bezeichnet (Rao/Molina, 2015, hier: Seite 2). Zur Berechnung präziser Schätzwerte für die entsprechenden Teilgesamtheiten können sogenannte Small-Area-Verfahren verwendet werden. Diese meist modellbasierten Ansätze beruhen darauf, dass zusätzliche Hilfsinformationen aus weiteren Teilgesamtheiten anhand eines vorher definierten statistischen Modells unterstützend einbezogen werden.

Die Anwendung von Small-Area-Verfahren auf Unternehmensdatensätze unterliegt jedoch verschiedenen Herausforderungen. Zunächst weisen Unternehmensdaten in der Regel schiefe Verteilungen mit einzelnen äußerst großen Ausreißern auf, deren Präsenz regionale Schätzwerte stark dominiert. Dies ist im Rahmen einer Small-Area-Schätzung zu berücksichtigen. Darüber hinaus sind Unternehmensregister eine potenzielle Quelle für Hilfsvariablen bei der Verwendung von Small-Area-Schätzverfahren. Zwischen der Zielgesamtheit, das heißt der Gesamtheit, die untersucht werden soll, und dem Register kommt es jedoch in der Praxis häufig zu Inkonsistenzen. Dabei können sich die Zielgesamtheit und das Register in der Zusammensetzung und der Anzahl der Unternehmen deutlich unterscheiden. Wenn nun die Daten des Unternehmensregisters als Hilfsvariablen für Small-Area-Verfahren verwendet werden, kann dies somit zu Problemen führen. Bleibt die fehlende

Konsistenz bei der Verwendung entsprechender Methoden unberücksichtigt, hat dies zur Folge, dass die Schätzer sogar unzuverlässiger sein können als klassische designgewichtete Schätzverfahren.

Aufbauend auf den beschriebenen Herausforderungen bestand das Ziel der Masterarbeit darin, das Potenzial zur Verbesserung der Schätzungen für Small Areas in Unternehmensstatistiken zu analysieren. In diesem Zusammenhang sollten insbesondere die erwähnten Inkonsistenzen berücksichtigt werden.

In Kapitel 2 werden daher zunächst verschiedene Arten und Ursachen von entsprechenden Inkonsistenzen genauer erläutert. Daraufhin wird eine Auswahl von Neugewichtungsansätzen für den Umgang mit Schichtwechslern vorgestellt. Das dritte Kapitel dieses Beitrags stellt den Battese-Harter-Fuller-Schätzer als Beispiel eines modellbasierten Small-Area-Verfahrens vor. Auf Basis der theoretischen Grundlage werden in Kapitel 4 schließlich die Umsetzung und zentrale Ergebnisse der Monte-Carlo-Simulationsstudie dargestellt. Der Beitrag schließt mit einem Fazit in Kapitel 5.

2

Neugewichtung bei Inkonsistenzen

Schichtwechsler

Die erwähnten Inkonsistenzen haben verschiedene Ursachen. Um das Stichprobendesign von Unternehmensstatistiken zu erstellen, wird für gewöhnlich das Statistische Unternehmensregister als Auswahlgesamtheit (auch: Auswahlgrundlage) verwendet. Dabei wird in der Regel jedoch ein Registerauszug verwendet, welcher im Hinblick auf den Berichtszeitraum der Statistik noch nicht aktuell ist (Statistische Ämter des Bundes und der Länder, 2012, hier: Seite 63). Darüber hinaus „erweist sich der von den Verwaltungen gemeldete [und im Unternehmensregister hinterlegte] Wirtschaftszweig häufig als fehlerhaft oder nicht mehr aktuell“ (Kless/Veldhues, 2008, hier: Seite 235) oder die Zuordnung der Unternehmen zu den Wirtschaftszweigen (WZ) ist nicht nach einheitlichen Vorgehensweisen erfolgt (Statistische Ämter des Bundes und der Länder, 2012, hier: Seite 63).

Ebenfalls können zeitliche Verzögerungen zwischen der Erstellung des Stichprobendesigns und der Datenerhebung selbst zu entsprechenden Inkonsistenzen führen. In diesem Kontext spricht man auch von Fehlern im Auswahlrahmen (englisch: frame errors).

Neben einer regionalen Schichtung umfasst das Stichprobendesign von Unternehmensstatistiken in der Regel eine Stratifizierung, das heißt die Zerlegung der Grundgesamtheit in überschneidungsfreie Teilgesamtheiten, nach Branchengruppen und Größenklassen, in die die Einheiten zum Beispiel entsprechend dem jährlichen Umsatz laut Register eingeteilt werden. Aufgrund der genannten Inkonsistenzen und der starken Dynamik von Unternehmensdaten resultieren jedoch Schichtwechsler. Dies sind Unternehmen, die einer anderen Designschicht zugeordnet worden wären, wenn die korrekten Designinformationen bereits zum Zeitpunkt des Stichprobendesigns bekannt gewesen wären. Die Problematik ist jedoch, dass diese Fehlzuordnungen erst bei der Befragung der entsprechenden Einheiten im Rahmen einer Erhebung entdeckt werden.

Im Folgenden wird zwischen branchenspezifischen Schichtwechslern und größenspezifischen Schichtwechslern unterschieden. Branchenspezifische Schichtwechsler sind Unternehmen, bei denen sich im Zuge der Befragung herausstellt, dass sie einer anderen Branche, das heißt einem anderen Wirtschaftszweig, zugehörig sind als dies im Register angegeben ist. Größenspezifische Schichtwechsler entstehen hingegen insbesondere durch die oben angesprochene mangelnde Aktualität des für das Stichprobendesign verwendeten Registers. Das sind Unternehmen, die im aktuellen Berichtsjahr nicht länger zu der ursprünglichen Größenklasse gehören. Dies geschieht bei starken Veränderungen der Variable, entsprechend der die Unternehmen in Größenklassen eingeteilt wurden.

Ein übliches Verfahren der direkten designgewichteten Schätzung ist der Schätzer nach Horvitz und Thompson (1952). Wie bereits in Manecke (2017a) erläutert, werden die erhobenen Kennzahlen bei diesem Verfahren auf die Grundgesamtheit hochgerechnet. Hierfür wird jedem befragten Unternehmen ein Designgewicht (auch: Hochrechnungsfaktor) zugewiesen, das der Kehrwert seiner jeweiligen Auswahlwahrscheinlichkeit ist. Die hochgerechneten Werte der befragten Unternehmen werden schließlich aufsummiert. Die Verwendung die-

ses Ansatzes wird im Kontext von größenspezifischen Schichtwechslern insbesondere dann problematisch, wenn ein Unternehmen ein deutliches Wachstum zu verzeichnen hat (Beaumont/Riest, 2009, hier: Seite 270). Bei der für Unternehmensstatistiken üblichen Allokation des Stichprobenumfangs wird vermeintlich kleinen Unternehmen, welche nur zu geringen Anteilen befragt werden, ein hohes Designgewicht zugeordnet. Wird nun ein fälschlicherweise für klein erachtetes Unternehmen befragt, so hat dies zur Folge, dass ein unerwartet hoher Merkmalswert irrtümlich mit einem hohen Designgewicht gewichtet wird. Dadurch werden die interessierenden Parameter der interessierenden Teilgesamtheiten deutlich überschätzt. Dies führt zu einer ineffizienten Schätzung.

Sind Schichtwechsler vorhanden, sind somit die Annahmen, unter denen die ursprünglichen Designgewichte innerhalb des Stichprobendesigns errechnet worden sind, nicht mehr erfüllt. Daher wurden im Rahmen der Masterarbeit sieben verschiedene Ansätze entwickelt, um die Designgewichte im Rahmen einer Neugewichtung anzupassen. Diese Ansätze lassen sich in fallzahlbasierte und modellbasierte Neugewichtungsverfahren unterteilen.

Bei fallzahlbasierten Verfahren wird das Designgewicht als die Anzahl der Einheiten in der Zielgesamtheit, die möglichst identische Variablenausprägungen wie die befragte Einheit haben, interpretiert. Unter Beachtung der Designschichten wird daher nach Berücksichtigung der Schichtwechsler versucht, die Anzahl ähnlicher Einheiten in der gleichen Auswertungsschicht so genau wie möglich zu rekonstruieren. Neben den Ergebnissen unter Verwendung der ursprünglichen und somit unveränderten Designgewichte (G_0) werden in Kapitel 4 die Resultate von zwei ausgewählten fallzahlbasierten Neugewichtungsansätzen (G_{F1} und G_{F2}) dargestellt. Während bei Ansatz G_{F1} die Unternehmen losgelöst von ihren Designgewichten die Schicht wechseln, wird im Rahmen von Ansatz G_{F2} davon ausgegangen, dass ein Schichtwechsel eines Unternehmens mit einem vorhandenen Muster einhergeht. Daher wird das jeweilige ursprüngliche Designgewicht von Schichtwechslern mit in die „neue“ Auswertungsschicht übertragen.

Bei den modellbasierten Neugewichtungsmethoden wird hingegen unter der Verwendung geeigneter Modelle der Zusammenhang der Designgewichte und der interes-

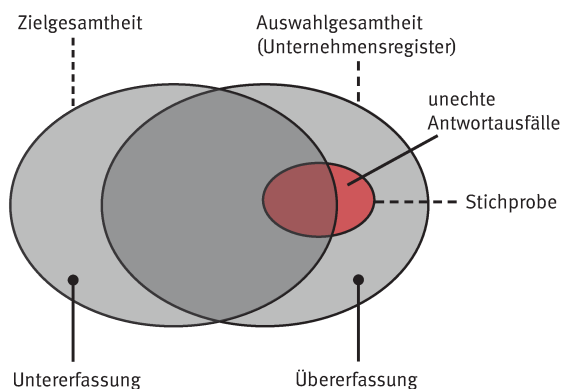
sierenden Variablen quantifiziert und zur Stabilisierung der Schätzung herangezogen (Beaumont, 2008; Beaumont/Rivest, 2009; Gershunskaya/Sverchkov, 2014). Dieser Ansatz wird auch als Gewichteglättung bezeichnet. In Kapitel 4 werden die Ergebnisse einer Gewichteglättung anhand eines inversen Exponentialmodells (G_M) beispielhaft für die modellbasierte Neugewichtung dargestellt.

Unechte Antwortausfälle

Neben Schichtwechslern kommt es bei der Abdeckung der Zielgesamtheit zu weiteren Inkonsistenzen. In Anlehnung an Schnell und andere (2011, hier: Seite 264) verdeutlicht [Grafik 1](#) das Verhältnis von der Zielgesamtheit, der Auswahlgesamtheit und der Stichprobe einer Erhebung bei vorhandenen Inkonsistenzen. Im Kontext der Jahrerhebung im Handel sowie in der Instandhaltung und Reparatur von Kfz des Jahres 2014 beinhaltet die Zielgesamtheit „alle Unternehmen mit Sitz in Deutschland, die [im Jahr 2014] ausschließlich oder überwiegend Handel betreiben“ (Statistisches Bundesamt, 2016, hier: Seite 1). Das heißt die Unternehmen sind den Wirtschaftszweigen (WZ) 45, 46 oder 47 zuzuordnen. Die Auswahlgesamtheit hingegen umfasst jene Unternehmen, die im Register, welches zur Erstellung des Stichprobendesigns verwendet worden ist, den Wirtschaftszweigen 45, 46 oder 47 zugeordnet sind.

Grafik 1

Verhältnis von Zielgesamtheit, Auswahlgesamtheit und Stichprobe einer Erhebung bei vorhandenen Inkonsistenzen



2019 - 01 - 0034

Unternehmen, die Teil der Zielgesamtheit, jedoch nicht Teil der verwendeten Auswahlgesamtheit sind, werden unter dem Begriff Untererfassung zusammengefasst. Hierzu gehören unter anderem Neuzugänge sowie Unternehmen, die im Register irrtümlich nicht den Wirtschaftszweigen 45, 46 oder 47 zugeordnet sind. Entsprechende Unternehmen werden nicht bei dem Design der Erhebung berücksichtigt und können somit auch nicht in die Stichprobe gelangen. Im Gegensatz dazu verursachen Unternehmen, die vor dem Berichtsjahr erloschen oder bereichsfremd sind, eine sogenannte Übererfassung, das heißt sie werden fälschlicherweise zur Auswahlgesamtheit gezählt.

Unternehmen, die Teil der Übererfassung sind, können Teil der Stichprobe werden. Da sie jedoch nicht zur Zielgesamtheit gehören und somit nicht als Ausfall zu werten sind, spricht man von unechten Antwortausfällen (Statistische Ämter des Bundes und der Länder, 2012, hier: Seite 54). Betroffene Unternehmen wurden in der Vergangenheit nicht in die entsprechende Statistik einbezogen und somit nicht mit ihren Designgewichten berücksichtigt¹. Infolgedessen führten die unechten Antwortausfälle bei der Auswertung der Handelsstatistik nicht zu einer Anhebung der Designgewichte der übrigen Unternehmen (ebenda). Es wird demnach indirekt angenommen, dass die Zahl der Unternehmen, die durch das jeweilige Designgewicht eines unechten Antwortausfalls „repräsentiert“ werden, ebenfalls Teil der Übererfassung ist. Im Gegenzug dazu fehlt jedoch die Kenntnis über die im Laufe des Berichtsjahres gegründeten Unternehmen sowie Unternehmen, die fälschlicherweise einem anderen Wirtschaftszweig zugeordnet sind. Diese Inkonsistenzen der Datenquellen führen dazu, dass die Statistiken die Realität unterzeichnen (Kless/Veldhues, 2008, hier: Seite 236; Statistische Ämter des Bundes und der Länder, 2012, hier: Seite 54). Dieser Ansatz, bei dem die Designgewichte unechter Antwortausfälle nicht berücksichtigt werden (Ansatz UA_1), wird in der Masterarbeit untersucht.

Aufgrund dieser starken Unterzeichnung wird zudem ein alternativer Ansatz (UA_2) im Hinblick auf verschiedene Schichtwechsel-Szenarien evaluiert. Dabei wer-

¹ Im Jahr 2017 wurden ab dem Berichtsjahr 2014 methodische Änderungen eingeführt. Seitdem werden unechte Antwortausfälle, die jedoch im Berichtsjahr nachweislich existierten, mit ihren Designgewichten berücksichtigt (Statistisches Bundesamt, 2017, hier: Seite 5).

den unechte Antwortausfälle wie echte Antwortausfälle behandelt und durch eine Erhöhung der Gewichte der übrigen befragten Einheiten kompensiert. Dabei wird indirekt vereinfachend angenommen, dass die Anzahl der Unternehmen je Schicht, die Teil der Übererfassung sind, im Gleichgewicht mit der Anzahl der Unternehmen steht, die als Neugründungen oder Bereichswechsler Teil der Untererfassung sind.

3

Small-Area-Schätzverfahren

Neben den Neugewichtungsansätzen wird zudem das Potenzial der Anwendung verschiedener Small-Area-Ansätze auf die Jahresehebung im Einzelhandel untersucht. Neben einer verbesserten Schätzgüte durch das Einbeziehen von Hilfsinformationen ermöglichen entsprechende modellbasierte Verfahren die Schätzung für Teilgesamtheiten, in denen keine Einheit befragt wurde. In der Masterarbeit wurden 17 verschiedene designbasierte, modellassistierte sowie modellbasierte Verfahren vergleichend gegenübergestellt. Für einen umfassenden Überblick über die unterschiedlichen Methoden sei an dieser Stelle auf Münnich und andere (2013) sowie Molina und Rao (2015) verwiesen.

In diesem Beitrag werden die Methodik und die Ergebnisse des sogenannten Battese-Harter-Fuller-Schätzers exemplarisch dargestellt (Battese und andere, 1988). Wie in Manecke (2017a) und Manecke (2017b) herausgestellt wird, handelt es sich dabei um ein Standardverfahren sogenannter Unit-Level-Modelle. Diese verwenden Hilfsinformationen auf Ebene der Erhebungseinheiten. Für alle D überschneidungsfreien Teilgesamtheiten der Zielgesamtheit mit dem Laufindex $d = 1, \dots, D$ soll der Totalwert τ_d der interessierenden Variablen geschätzt werden. Es sei y_{dk} die interessierende Variable von Unternehmen k in der Teilgesamtheit d und \mathbf{x}_{dk} der Vektor mit bekannten Hilfsvariablen dieser Einheit. Zudem ist N_d die Anzahl der Unternehmen, die zu Teilgesamtheit d gehören. Das dem Battese-Harter-Fuller-Schätzer zugrunde liegende und für die Grundgesamtheit angenommene Modell ist eine Sonderform des allgemeinen gemischten linearen Regressionsmodells und gegeben durch

$$y_{dk} = \mathbf{x}_{dk}^T \boldsymbol{\beta} + v_d + e_{dk}, \quad d = 1, \dots, D, \quad k = 1, \dots, N_d$$

$$v_d \stackrel{iid}{\sim} N(0, \sigma_v^2)$$

$$e_{dk} \stackrel{iid}{\sim} N(0, \sigma_e^2)$$

Hierbei ist $\mathbf{x}_{dk}^T \boldsymbol{\beta}$ der Regressionsterm mit dem Vektor der Regressionskoeffizienten $\boldsymbol{\beta}$, welcher über alle Teilgesamtheiten und Einheiten die fixen Effekte, das heißt den Zusammenhang zwischen der interessierenden Variable y_{dk} und den Hilfsvariablen \mathbf{x}_{dk} , erfasst. Der sogenannte Zufallseffekt v_d berücksichtigt Schwankungen zwischen den Teilgesamtheiten, die nicht durch den fixen Effekt des Regressionsterms erklärt werden können. Zudem beschreibt die Störgröße e_{dk} die individuellen Stichprobenfehler. Dabei haben v_d und e_{dk} jeweils den Erwartungswert 0 und die Varianzen σ_v^2 beziehungsweise σ_e^2 , wobei angenommen wird, dass beide Größen unabhängig voneinander sind. Die Varianz der Zufallseffekte σ_v^2 wird auch als Modellvarianz bezeichnet, da sie die Varianz zwischen den Teilgesamtheiten angibt, die nicht durch die fixe Komponente des Modells erklärt werden kann.

Unter der Annahme, dass das obige gemischte Regressionsmodell ebenfalls für die Stichprobe gilt, wird der Mittelwert der interessierenden Variable je Teilgesamtheit μ_d geschätzt durch den sogenannten Empirical Best Linear Unbiased Predictor (EBLUP) nach Battese und andere (1988), formal gegeben durch

$$\hat{\mu}_d^{BHF} = \bar{\mathbf{x}}_d^T \hat{\boldsymbol{\beta}} + \hat{v}_d \text{ mit}$$

$$\hat{v}_d = \hat{\gamma}_d (\bar{y}_d - \bar{\mathbf{x}}_d^T \hat{\boldsymbol{\beta}}) \text{ und}$$

$$\hat{\gamma}_d = \frac{\hat{\sigma}_v^2}{\hat{\sigma}_v^2 + \hat{\sigma}_e^2 / n_d}.$$

Hierbei bezeichnet $\bar{\mathbf{x}}_d$ den bekannten Durchschnitt der verwendeten Hilfsvariablen in der gesamten Teilgesamtheit d . Zudem sind $\bar{\mathbf{x}}_d$ und \bar{y}_d die Stichprobendurchschnitte der Hilfsvariablen beziehungsweise der interessierenden Variablen je Teilgesamtheit d . Darüber hinaus ist n_d die entsprechende Größe der Stichprobe in d . Da die Modellvarianz σ_v^2 und die Varianz des Stichprobenfehlers σ_e^2 in der Realität unbekannt sind, werden diese zur Ermittlung des EBLUPs geschätzt. Zur Schätzung der Varianzkomponenten gibt es verschiedene Verfahren. Für die im Folgenden dargestellte Unit-Level-

Schätzung wurde die Restricted-Maximum-Likelihood-Methode verwendet. Der Battese-Harter-Fuller-Schätzer für Teilgesamtheit d ist dementsprechend eine Kombination aus einem synthetischen Schätzer $\bar{X}_d^T \hat{\beta}$, der sich aus der fixen Modellkomponente bildet, sowie dem jeweiligen geschätzten Zufallseffekt \hat{v}_d . Durch die Erfassung der Zusammenhänge der Hilfsvariablen und der interessierenden Variablen über die innerhalb der Stichprobe befragten Einheiten werden also die entsprechenden Parameter anhand des gemischten linearen Regressionsmodells geschätzt und ermöglichen so in Kombination mit den Populationsdurchschnitten der Hilfsvariablen eine Berechnung von Schätzwerten für die Grundgesamtheit.

Der EBLUP lässt sich umformen zu der zusammengesetzten Schätzfunktion

$$\hat{\mu}_d^{BHF} = \hat{\gamma}_d (\bar{y}_d + (\bar{X}_d - \bar{x}_d)^T \hat{\beta}) + (1 - \hat{\gamma}_d) \bar{X}_d^T \hat{\beta}.$$

Hier ist zu erkennen, dass der Battese-Harter-Fuller-Schätzer ein gewichtetes Mittel aus dem Stichprobenregressionsschätzer $\bar{y}_d + (\bar{X}_d - \bar{x}_d)^T \hat{\beta}$ und dem synthetischen Schätzer $\bar{X}_d^T \hat{\beta}$ ist. Der Gewichtungsfaktor $\hat{\gamma}_d$ gibt je Teilgesamtheit den Anteil der Modellvarianz an der Gesamtvarianz an und legt fest, wie viel Gewicht auf die jeweiligen Komponenten gelegt wird. Bei einer sehr hohen Modellvarianz $\hat{\sigma}_v^2$ oder einem großen Stichprobenumfang n_d wird entsprechend viel Vertrauen in den direkten Stichprobenregressionsschätzer gelegt. Im Gegenzug dazu nähert sich der EBLUP bei einer geringen Modellvarianz oder einem geringen Teilstichprobenumfang tendenziell der synthetischen Komponente an, welche die Variablenausprägungen der Stichprobeneinheiten aus Teilgesamtheit d nicht berücksichtigt. Entsprechend besteht der EBLUP bei Teilgesamtheiten, in denen keine Einheit befragt worden ist ($n_d = 0$), vollständig aus dem synthetischen Schätzer, sodass hier $\mu_d^{BHF} = \bar{X}_d^T \hat{\beta}$ gilt.

Zur Verringerung des Einflusses von äußerst großen Unternehmen ist bei dem im Folgenden vorgestellten Battese-Harter-Fuller-Schätzansatz eine Logarithmierung der interessierenden Variablen y_{dk} sowie der Hilfsvariablen x_{dk} durchgeführt worden. Die Modellschätzung erfolgt dabei anhand der logarithmierten Werte. Der resultierende EBLUP schätzt schließlich den Mittelwert der logarithmierten Zielvariablen je Teilgesamtheit. Anhand einer Rücktransformation des EBLUPs durch die Exponentialfunktion entsteht schließlich der Schätzer

für den Mittelwert der interessierenden Variablen je Teilgesamtheit $\hat{\mu}_d^{BHF}$. Der endgültige Prädiktor des Totalwertes der interessierenden Variablen $\hat{\tau}_d^{BHF}$ ist schließlich das Produkt aus der bekannten Anzahl der Unternehmen aus Teilgesamtheit d in der Grundgesamtheit, N_d , und dem hergeleiteten Mittelwertschätzer, und ist somit gegeben durch $\hat{\tau}_d^{BHF} = N_d \hat{\mu}_d^{BHF}$.

4

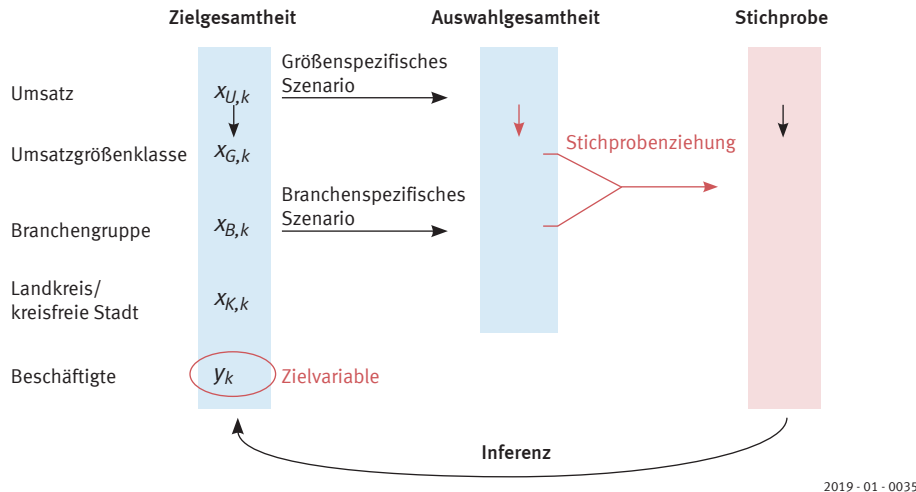
Simulationsstudie

Die theoretisch fundierten Ansätze werden schließlich in einer Simulationsstudie angewendet und miteinander verglichen. Die Simulationsstudie wird in starker Anlehnung an die Jahreserhebung im Einzelhandel des Bundeslandes Hessen (Hessisches Statistisches Landesamt, 2017) durchgeführt. Hierfür wird ein anonymisierter, synthetischer Datensatz der Unternehmen, welche im Auszug des hessischen Unternehmensregisters des Jahres 2014 dem WZ 47 (Einzelhandel ohne Handel mit Kfz) zugeordnet waren, verwendet. Das Ziel der Simulationsstudie ist es, die Gesamtzahl der sozialversicherungspflichtig Beschäftigten je Teilgesamtheit zu schätzen. Diese Teilgesamtheiten definieren sich dabei als Kombination aus zehn zur Schichtung verwendeten Branchengruppen des Einzelhandels (die WZ-Dreisteller 47.1 bis 47.9; der WZ-Dreisteller 47.7 ist zudem in zwei Unter-Branchengruppen unterteilt) und den 26 Landkreisen beziehungsweise kreisfreien Städten des Bundeslandes Hessen. Es resultieren somit 260 interessierende Teilgesamtheiten.

Der Aufbau der Simulation mit 2 000 Monte-Carlo-Replikationen ist in [Grafik 2](#) dargestellt. Der vorliegende synthetische Datensatz wird als wahre Zielgesamtheit definiert, die in der Praxis in der Regel nicht in dem Umfang bekannt ist. Basierend auf dieser Zielgesamtheit wird in jeder Replikation eine Auswahlgesamtheit generiert, die als in der Praxis vorliegendes Register angenommen wird. Anhand der wahren Umsatzwerte $x_{U,k}$ werden dann unter Verwendung verschiedener größenspezifischer Schichtwechsel-Szenarien veraltete Umsatzwerte $\tilde{x}_{U,k}$ generiert. Darüber hinaus werden im Rahmen unterschiedlicher branchenspezifischer Schichtwechsel-Szenarien die WZ-Zuordnungen laut Register $\tilde{x}_{B,k}$ erstellt. Diese Auswahlgesamtheit enthält somit sowohl größen- als auch branchenspezifische Schicht-

Grafik 2

Definition der Auswahlgesamtheit und der Stichprobenziehung in der Simulationsstudie



wechsler. Insgesamt wurden in der Masterarbeit Kombinationen aus sieben branchenspezifischen und sechs größenspezifischen Szenarien generiert. Somit kann untersucht werden, inwieweit die entwickelten Ansätze eine Verbesserung gegenüber dem klassischen designgewichteten Horvitz-Thompson-Schätzer in verschiedenen realistischen Inkonsistenz-Szenarien erreichen können. Schließlich wird die Stichprobe aus dieser Auswahlgesamtheit unter Verwendung der tatsächlichen Schichtmerkmale, Grenzen der Umsatzgrößenklassen und schichtspezifischen Auswahlätze der hessischen Einzelhandelerhebung des Jahres 2014 gezogen, um den Schichtungs-, Ziehungs- und Schätzprozess dieser Erhebung so realistisch wie möglich zu replizieren.

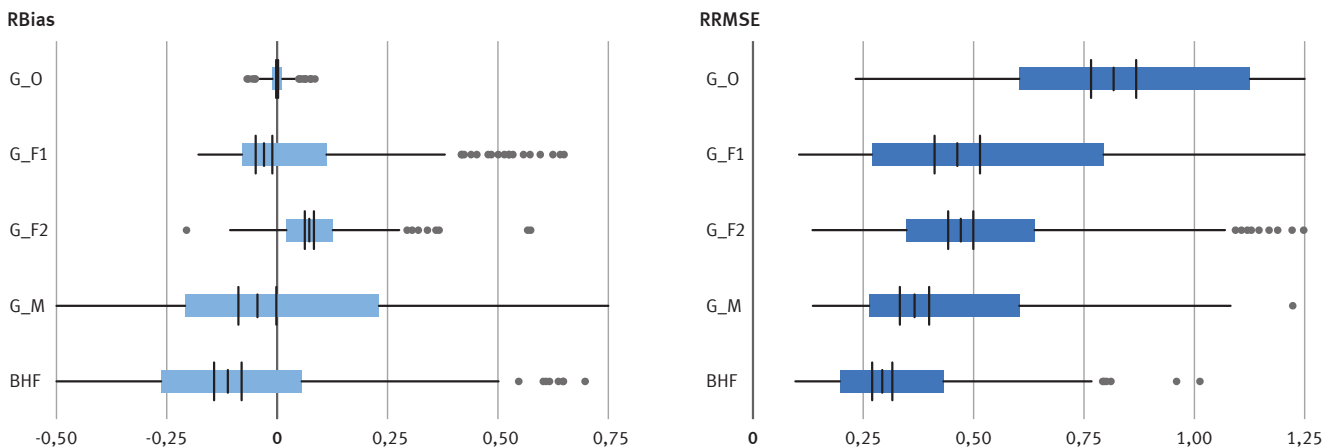
Es wird davon ausgegangen, dass die interessierende Variable y_k , das heißt die Anzahl sozialversicherungspflichtig Beschäftigter, nur für die Einheiten der Stichprobe bekannt ist. Mithilfe der Stichprobe soll somit die Gesamtzahl sozialversicherungspflichtig Beschäftigter in jeder Teilgesamtheit der Zielgesamtheit geschätzt werden. Dies wird anhand der entwickelten Neugewichtungsansätze zur Gewichtung des Horvitz-Thompson-Schätzers sowie verschiedener Small-Area-Methoden durchgeführt. In der Simulation wird der dargestellte Prozess je Kombination aus Schätzverfahren und Szenario insgesamt 2000-mal repliziert. Um die Schätzgüte der Schätz- sowie Gewichtungsansätze in den verschiedenen branchenspezifischen und größenspezifischen Szenarien zu beurteilen, werden je Kombination die

2000 Schätzwerte je Teilgesamtheit mit dem wahren Totalwert der Zielvariablen in der Teilgesamtheit der Zielgesamtheit verglichen. Als Maß für die Verzerrung des Schätzers dient der Relative Bias (RBias). Zur Beurteilung der Präzision wird die relative Wurzel aus der mittleren quadratischen Abweichung (RRMSE: Relative Root Mean Squared Error) verwendet.

➤ Grafik 3 zeigt die Simulationsergebnisse für die ausgewählten Neugewichtungsansätze und den Battese-Harter-Fuller-Schätzer für eine Kombination aus realistischen Schichtwechsel-Szenarien. Dabei sind in jedem Boxplot die Präzisionsmaße der 260 interessierenden Teilgesamtheiten abgebildet. Es wird ersichtlich, dass insbesondere der Horvitz-Thompson-Schätzer bei der Beibehaltung der ursprünglichen Designgewichte (G_O) zwar unverzerrt, jedoch auch im Vergleich zu den anderen Schätzansätzen sehr ineffizient ist, da er hohe RRMSE-Werte aufweist. Die Schätzung unter Verwendung der untersuchten fallzahl- und modellbasierten Neugewichtungsansätze (G_{F1} , G_{F2} und G_M) weist dahingegen Verzerrungen auf. Diese sind insbesondere bei dem modellbasierten Neugewichtungsansatz (G_M) in einem Großteil der Teilgesamtheiten sehr ausgeprägt. Mit Blick auf den RRMSE kann bei der Schätzung anhand der im Rahmen der Neugewichtung angepassten Gewichte (G_{F1} , G_{F2} und G_M) jedoch eine signifikante Verbesserung gegenüber der Schätzung auf Basis der ursprünglichen Designgewichte (G_O) beobachtet werden. Hier ist insbesondere der modellbasierte Neu-

Grafik 3

Ergebnisse der Simulationsstudie: Relativer Bias und RRMSE für die Gewichtungsmethoden und den Battese-Harter-Fuller-Schätzer



RBias: Relativer Bias; RRMSE: Relative Root Mean Squared Error (relative Wurzel aus der mittleren quadratischen Abweichung); G_O: Horvitz-Thompson-Schätzer bei der Beibehaltung der ursprünglichen Designgewichte; G_F1, G_F2 und G_M: untersuchte fallzahl- und modellbasierte Neugewichtungsansätze; BHF: Battese-Harter-Fuller-Schätzer.

2019 - 01 - 0036

gewichtungsansatz (G_M) sehr robust gegenüber starken Schichtwechsel-Strukturen.

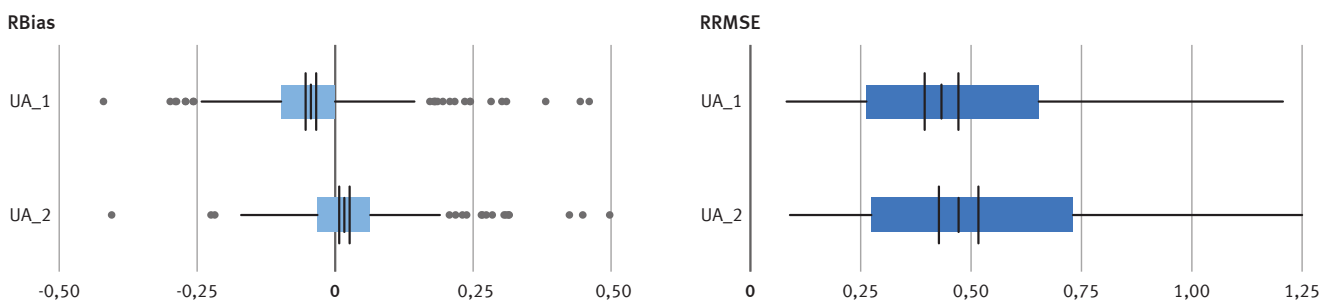
Der Battese-Harter-Fuller-Schätzer (BHF) zieht den Umsatz laut Register als Hilfsvariable in die Schätzung der Anzahl sozialversicherungspflichtig Beschäftigter unterstützend mit ein. Vergleicht man die Ergebnisse der Schätzung mit dem einfachen designgewichteten Horvitz-Thompson-Schätzer, so ist zu erkennen, dass durch die Einbeziehung der Hilfsinformationen eine deutliche zusätzliche Verbesserung der Schätzgüte erzielt werden

kann. Durch die Logarithmierung der Ziel- und Hilfsvariablen kann zudem der starke Einfluss von Ausreißern auf die Modellparameterschätzung reduziert werden. So ist der Schätzer auch bei deutlich veralteten Hilfsvariablen im Rahmen starker Schichtwechsel-Szenarien auffallend robust.

Darüber hinaus sind die in Kapitel 2 vorgestellten Ansätze im Umgang mit unechten Antwortausfällen, das heißt Einheiten der Stichprobe, die jedoch nicht zu der Zielgesamtheit gehören, simulativ untersucht worden.

Grafik 4

Ergebnisse der Simulationsstudie: Vergleich der Ansätze im Umgang mit unechten Antwortausfällen



RBias: Relativer Bias; RRMSE: Relative Root Mean Squared Error (relative Wurzel aus der mittleren quadratischen Abweichung); UA_1: Designgewichte unechter Antwortausfälle werden nicht berücksichtigt; UA_2: Umverteilung der Designgewichte der unechten Antwortausfälle im Rahmen des alternativen Ansatzes.

2019 - 01 - 0037

Für diese Fragestellung sind im Rahmen der Simulation in entsprechenden Szenarien ebenfalls eine Über- und Untererfassung der Auswahlgesamtheit generiert worden. Die Ergebnisse sind in [Grafik 4](#) zu sehen. Im Hinblick auf die Verfahrensweise, bei welcher die Designgewichte unechter Antwortausfälle nicht berücksichtigt werden (Ansatz UA_1), kann die Literatur bestätigt werden: Die Totalwerte werden bei dieser Verfahrensweise sichtbar unterschätzt. Dieser Unterschätzung kann durch eine Umverteilung der Designgewichte der unechten Antwortausfälle im Rahmen des alternativen Ansatzes (UA_2) entgegengewirkt werden. Dieser Ansatz führt unter den Schichtwechsel-Annahmen der Simulationsstudie zu einer signifikanten Reduzierung der bisherigen Unterschätzung. Allerdings hat diese Umverteilung ebenfalls eine Verringerung der Schätzeffizienz zur Folge, was an dem erhöhten RRMSE zu erkennen ist.

sität Trier mit dem Hessischen Statistischen Landesamt entstanden. In diesem Zusammenhang möchte ich mich sehr herzlich beim Hessischen Statistischen Landesamt, insbesondere bei Patrik Vollmer und Dr. Peter Gottfried, für die sehr gute Zusammenarbeit bedanken. Ein besonderer Dank gilt zudem Prof. Dr. Ralf Münnich und Dr. Jan Pablo Burgard für die Unterstützung und die sehr gute fachliche Betreuung der Arbeit. [\[1\]](#)

5

Fazit

Es ist somit abschließend festzuhalten, dass sowohl bei der Analyse von Small-Area-Verfahren als auch beim Vergleich von Gewichtungsansätzen ein deutlicher Zielkonflikt zwischen der Verzerrung und der Varianz der Schätzung besteht. Der klassische Horvitz-Thompson-Schätzer unter Beibehaltung der ursprünglichen Designgewichte führt selbst bei starken Inkonsistenzen zu unverzerrten Schätzungen. Diese gehen jedoch in der Regel mit einer inakzeptabel hohen Schätzvarianz einher. Obwohl Neugewichtungsansätze sowie modellbasierte Small-Area-Verfahren wie der Battese-Harter-Fuller-Schätzer zu einer Verzerrung der Schätzung führen, kann in Bezug auf die Schätzvarianz eine deutliche Erhöhung der Schätzgüte im Vergleich zu dem designgewichteten Horvitz-Thompson-Schätzer festgestellt werden. Trotz vorhandener Inkonsistenzen und der Verwendung ungenauer Hilfsvariablen kann dadurch eine stabilisierte Schätzung erzielt werden.

Danksagung

Die Masterarbeit ist im Rahmen des European Master in Official Statistics (EMOS) aus einer Kooperation der Professur für Wirtschafts- und Sozialstatistik der Univer-

LITERATURVERZEICHNIS

Battese, George E./Harter, Rachel M./Fuller, Wayne A. *An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data*. In: Journal of the American Statistical Association. Jahrgang 83. Ausgabe 401. März 1988, Seite 28 ff.

Beaumont, Jean-François. *A new approach to weighting and inference in sample surveys*. In: Biometrika. Jahrgang 95. Ausgabe 3. September 2008, Seite 539 ff.

Beaumont, Jean-François/Rivest, Louis-Paul. *Dealing with Outliers in Survey Data*. In: Pfeffermann, Danny/Rao, Calyampudi R. (Herausgeber). Handbook of Statistics. Jahrgang 29. Teil A. 2009, Seite 247 ff.

Gershunskaya, Julie/Sverchkov, Michael. *On Weight Smoothing in the Current Employment Statistics Survey*. In: JSM 2014 – Survey Research Methods Section. Bureau of Labor Statistics. 2014, Seite 1139 ff.

Hessisches Statistisches Landesamt. [Strukturdaten des Einzelhandels in Hessen im Jahr 2014. Ergebnisse der Jahrerhebung](#). In: Statistische Berichte. 2017.

Horvitz, Daniel G./Thompson, Donovan J. *A Generalization of Sampling Without Replacement From a Finite Universe*. In: Journal of the American Statistical Association. Jahrgang 47. Ausgabe 260/1952, Seite 663 ff.

Kless, Sascha/Veldhues, Bernhard. [Ausgewählte Ergebnisse für kleine und mittlere Unternehmen in Deutschland 2005](#). In: Wirtschaft und Statistik. Ausgabe 3/2008, Seite 225 ff.

Manecke, Julia. [Die Regionalisierung der Hessischen Jahrerhebung im Einzelhandel anhand von Small-Area-Verfahren](#). In: StaWi – Staat und Wirtschaft in Hessen. Ausgabe 1/2017a, Seite 29 ff.

Manecke, Julia. [Regionale Auswertung von Unternehmensstatistiken: Methoden und Anwendungen im Kontext der Small Area-Statistik](#). Unveröffentlichte Masterarbeit. 2017b.

Münnich, Ralf/Burgard, Jan Pablo/Vogt, Martin. *Small Area-Statistik: Methoden und Anwendungen*. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 6. Ausgabe 3–4/2013, Seite 149 ff.

Rao, J.N.K./Molina, Isabel. *Small Area Estimation*. 2. Auflage. Hoboken/New Jersey 2015.

Schnell, Rainer/Hill, Paul B./Esser, Elke. *Methoden der empirischen Sozialforschung*. 9. Auflage. München 2011.

Statistische Ämter des Bundes und der Länder. [Reform der Unternehmensstatistik, Endbericht](#). 2012.

Statistisches Bundesamt. [Jahrerhebung im Handel sowie in der Instandhaltung und Reparatur von Kfz 2014](#). Qualitätsbericht. 2016.

Statistisches Bundesamt. [Jahrerhebung im Handel sowie in der Instandhaltung und Reparatur von Kfz 2015](#). Qualitätsbericht. 2017.

Herausgeber

Statistisches Bundesamt (Destatis), Wiesbaden

Schriftleitung

Dr. Sabine Bechtold

Redaktionsleitung: Juliane Gude

Redaktion: Ellen Römer

Ihr Kontakt zu uns

www.destatis.de/kontakt

Erscheinungsfolge

zweimonatlich, erschienen im Februar 2019

Das Archiv aller Ausgaben ab Januar 2001 finden Sie unter www.destatis.de/publikationen

Print

Einzelpreis: EUR 19,- (zzgl. Versand)

Jahresbezugspreis: EUR 114,- (zzgl. Versand)

Bestellnummer: 1010200-19001-1

ISSN 0043-6143

ISBN 978-3-8246-1080-8

Download (PDF)

Artikelnummer: 1010200-19001-4, ISSN 1619-2907

Vertriebspartner

IBRo Versandservice GmbH

Bereich Statistisches Bundesamt

Kastanienweg 1

D-18184 Roggentin

Telefon: +49 (0) 382 04 / 6 65 43

Telefax: +49 (0) 382 04 / 6 69 19

destatis@ibro.de

Papier: Design Offset, FSC-zertifiziert

© Statistisches Bundesamt (Destatis), 2019

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.