

LUDWIGS-MAXIMILIAN UNIVERSITÄT
MÜNCHEN

SPEZIELLE THEMEN DER SOZIAL- UND
WIRTSCHAFTSSTATISTIK

Small Area-Schätzung

Autor:
Frederik
LUDWIGS

Betreuung
Prof. Dr. Thomas
AUGUSTIN



9. Juni 2016

Abstract

In diesem Bericht werden Verfahren vorgestellt, die es erlauben für regionale oder inhaltliche Subgruppen reliable Schätzungen abzugeben. Diese Verfahren werden unter dem Begriff Small Area-Schätzung zusammengefasst. Solche Verfahren ermöglichen valide Schätzungen für Subgruppen, selbst wenn der Stichprobenumfang in der entsprechenden Subgruppe für klassische Schätzverfahren zu gering ist. Dabei werden, anders als bei klassischen Schätzverfahren, Hilfsinformationen aus übergeordneten Gebieten bzw. Populationen genommen um so den effektiven Stichprobenumfang zu vergrößern.

Zuerst werden klassischen Schätzverfahren, wie beispielsweise dem Horvitz-Thompson Schätzer und dem verallgemeinerten Regressionsschätzer, vorgestellt. Neben den allgemeinen theoretischen Grundlagen der Small Area-Schätzung, wird auch ein praktisches Anwendungsbeispiel aus der amtlichen Statistik vorgestellt. Anhand dieses Anwendungsbeispiels werden die verschiedenen Verfahren miteinander verglichen und deren Genauigkeit bewertet. Dabei liefern die Schätzungen mit Small Area-Methoden im Vergleich zu den klassischen Schätzverfahren in diesem Anwendungsbeispiel wesentlich bessere Ergebnisse.

Inhaltsverzeichnis

1	Einleitung	3
2	Grundlegende Informationen zum Anwendungsbeispiel	4
3	Grundlegende Schätzverfahren	5
3.1	Designbasierte Ansätze	5
3.1.1	Horvitz-Thompson-Schätzer	5
3.1.2	Verallgemeinerter Regressionsschätzer	7
4	Grundlagen der Small Area-Schätzung	9
4.1	Begriffliche Grundlagen	9
4.2	Designbasierte Area-Schätzung	10
4.3	Synthetische Schätzungen	11
4.4	Zusammengesetzte Schätzer	13
5	Spezielle Small Area-Schätzmethoden	16
5.1	Battese-Harter-Fuller-Schätzer	17
5.2	Fay-Herriot Schätzer	17
6	Vergleich der verschiedenen Schätzverfahren nach [Dieterle, 2010]	19
6.1	Anwendung der Schätzverfahren im Beispiel der Schweinebestandschätzung	19
6.2	Verfahren zum Vergleich der verschiedenen Schätzverfahren nach [Dieterle, 2010]	20
6.3	Ergebnisse der Vergleiche	21
7	Fazit und Ausblick	27

1 Einleitung

Ob nun für die Armutsmessung, den Zensus oder eine tiefere Untergliederung von Indikatoren, sowohl die Forschung, wie auch Wirtschaft und Politik benötigen immer genauere regionalisierte Informationen. So ist es nicht verwunderlich, dass in den letzten Jahren die Nachfrage nach eben diesen stark gestiegen ist. Doch nicht nur die Nachfrage nach regionalen Informationen steigt, sondern die Nachfrage nach Informationen von Subgruppen, seien es geografische oder inhaltliche, insgesamt. Wie wichtig dies ist, lässt sich in Europa an der Finanzierung von Forschungsprogrammen gut erkennen, so wurden bei dem fünften, sechsten und siebten Forschungsrahmenprogrammen der Europäischen Kommission immer wieder Projekte gefördert, die sich eben mit dieser Thematik befassen. „Dem [...] zunehmenden Bedarf an detaillierten Informationen stehen jedoch überwiegend unveränderte oder niedrigere Budgets für die Datenerhebung gegenüber“ [Münnich et al., 2013, S. 151]. Um eben dieser Widersprüchlichkeit von höheren Nachfragen nach qualitativ hochwertigen und genauen Daten und den sinkenden Budgets entgegen zu wirken, werden Verfahren der *Small Area-Schätzung* benötigt. Solche Methoden verwenden dafür Modelle, die neben der interessierenden Teilgesamtheit noch weitere, verfügbare Informationen aus benachbarten Gebieten oder ähnlichen Gruppen in den Schätzungen berücksichtigen können. Dies hat zur Folge, dass bei Gebieten bzw. Subgruppen mit kleinen Stichprobenumfängen besser geschätzt werden kann und damit die Schätzungen mit dem Verfahren der Small Area-Schätzung im Vergleich zu klassischen Schätzverfahren wesentlich genauer sind.

Im folgenden sollen die grundlegenden Methoden der Small Area-Schätzung erläutert werden und anhand eines Anwendungsgebietes in der amtlichen Statistik veranschaulicht werden. Als Anwendungsbeispiel soll die Schätzung von Schweinebeständen dienen.

2 Grundlegende Informationen zum Anwendungsbeispiel

2012 veröffentlichte das statistische Bundesamt Wiesbaden einen Bericht zu dem Thema der Small Area-Schätzmethoden. In diesem werden verschiedene Schätzverfahren, sowie deren Vor- und Nachteile, anhand des Beispiels der Schweinebestandsschätzung dargestellt. Die Bestände von Schweinen sind deshalb so wichtig, da sich Deutschland „im Rahmen verschiedener internationaler Verträge wie dem Genfer Luftreinhalteabkommen“ [Dieterle, 2010, S. 1212] dazu verpflichtet hat über die Emissionen klimarelevanter Gase, welche eben auch Schweine produzieren, zu berichten. Seit 1999 finden Totalerhebungen nicht mehr alle zwei Jahre statt, sowie es davor üblich war, sondern wesentlich seltener. Für die Zeiträume zwischen den Totalerhebungen werden die Bestände auf Länderebene, allerdings nicht auf Kreisebene, zwei mal jährlich geschätzt. Da die Stichprobe eigentlich für eine valide Schätzung auf Länderebene und nicht für valide Schätzungen auf Kreisebene ausgelegt ist, werden Small Area-Schätzmethoden benötigt, um valide Ergebnisse auf Kreisebene schätzen zu können. Da für das Jahr 2007 sowohl Stichproben- als auch Totalmaterial vorliegt ist die Quantifizierung der Fehler der Schätzverfahren in diesem Fall möglich, sodass auch ein Vergleich der verschiedenen Schätzverfahren möglich ist. Dabei war das Vorgehen wie folgt: Zuerst wurden die Schweinebestände auf Kreisebene im Jahr 2007 anhand verschiedener Verfahren auf Basis der Stichprobe aus dem Jahr 2007 geschätzt und anschließend mit den wahren Werten verglichen. Dies dient vor allem zur Veranschaulichung der vorgestellten Verfahren und soll Unterschiede, sowie Vor- und Nachteile der verschiedenen Verfahren deutlich machen.

3 Grundlegende Schätzverfahren

3.1 Designbasierte Ansätze

In der amtlichen Statistik sind bei vielen Anwendungen vor allem Total-, Durchschnitts-, oder Anteilwerte, die sich gut in Tabellen veröffentlichen lassen, von Interesse. Dabei bilden designbasierte Ansätze die Grundlage moderner Stichprobentheorie.

Es wird zunächst von einer endlichen Grundgesamtheit Ω ausgegangen, welche aus N Elementen besteht. Die Ausprägungen des interessierenden Merkmales werden hierbei mit y_i ($i = \{1, 2, \dots, N\}$) bezeichnet. Aus dieser Grundgesamtheit Ω wird eine Stichprobe φ , welche durch „sukzessive Auswahl von Elementen“ [Münnich et al., 2013, S. 153] entsteht, vom Umfang n gezogen. Die Wahrscheinlichkeit, dass das Element i in die Stichprobe φ aufgenommen wird, wird als *Inklusionswahrscheinlichkeit erster Ordnung* π_i bezeichnet und berechnet sich im Falle einer einfachen Stichprobe über $\pi_i = \frac{n}{N}$. Die *Inklusionswahrscheinlichkeit zweiter Ordnung* gibt die Wahrscheinlichkeit dafür an, dass sowohl Element i , als auch Element j in die Stichprobe φ gelangen und wird als $\pi_{i,j}$ bezeichnet. Das Design-Gewicht ω_i berechnet sich über $\omega_i = \frac{1}{\pi_i}$. Hierbei ist das Design-Gewicht ω_i , umso höher, umso kleiner die Inklusionswahrscheinlichkeit erster Ordnung π_i ist.

Angewandt auf unser Beispiel, bezeichnet Ω die Menge aller Schweinehaltenden Betriebe, y_i den Schweinebestand der Stichprobeneinheit i und π_i die Wahrscheinlichkeit, dass ein Schweinehaltender Betrieb in die Stichprobe φ gelangt.

3.1.1 Horvitz-Thompson-Schätzer

Die zwei Statistiker D.G. Horvitz und D.J. Thompson stellten im Jahre 1952 in ihrer Arbeit *A generalization of sampling without replacement from a finite universe* einen Schätzer vor, der heutzutage als Horvitz-Thompson-Schätzer bekannt ist. Dieser Schätzer eignet sich zur Schätzung von Total- und Mittelwerten und ist in allen Fällen anwendbar, bei denen kein Element öfter als einmal in die Stichprobe gelangen kann. Nach [Kauermann and Küchenhoff, 2010, S. 95] ist die Idee dieses Schätzers dabei, anhand der Inklusionswahrscheinlichkeiten π_i einen unverzerrten Schätzer zu konstruieren.

Ist der Totalwert τ , die Summe über alle Beobachtungswerte, des Merkmals Y von Interesse, so wird dieser über

$$\hat{\tau}_Y^{HT} = \sum_{i \in \varphi} \frac{y_i}{\pi_i} = \sum_{i \in \varphi} y_i \omega_i \quad (1)$$

geschätzt. Der Mittelwert der Grundgesamtheit wird über

$$\hat{\mu}_Y^{HT} = \frac{1}{\hat{N}} \sum_{i \in \varphi} \frac{y_i}{\pi_i} = \sum_{i \in \varphi} y_i \omega_i \quad (2)$$

geschätzt, wobei \hat{N} über $\sum_{i \in \varphi} \frac{1}{\pi_i}$ geschätzt wird. „Die Verwendung der geschätzten anstatt der wahren Umfänge im Nenner führt bei Stichprobendesigns mit variablen Umfängen oft zu besseren Schätzungen, da hier sehr große und kleine Stichprobenumfänge in der Schätzung kompensiert werden können“ [Münnich et al., 2013]. Bei dem Horvitz-Thompson Schätzverfahren werden also die Merkmalsausprägungen der Stichprobe entsprechend gewichtet, sodass eine Merkmalsausprägung umso stärker ins Gewicht fällt, umso geringer die Inklusionswahrscheinlichkeit erster Ordnung π_i ist.

Hierbei lässt sich die Varianz des Schätzers für den Totalwert $\hat{\tau}_Y^{HT}$ über

$$\hat{V}(\hat{\tau}_Y^{HT}) = \sum_{i \in \varphi} \frac{(1 - \pi_i)}{\pi_i^2} y_i^2 + 2 \sum_{i \in \varphi} \sum_{j \in \varphi} \left(1 - \frac{\pi_i \pi_j}{\pi_{i,j}}\right) \frac{y_i}{\pi_i} \frac{y_j}{\pi_j} \quad (3)$$

schätzen und die Varianz für den geschätzten Mittelwert über

$$\hat{V}(\hat{\mu}_Y^{HT}) = \frac{1}{N^2} \left[\sum_{i \in \varphi} \frac{(1 - \pi_i)}{\pi_i^2} y_i^2 + \sum_{i \in \varphi} \sum_{j \in \varphi} \frac{\pi_{i,j} - \pi_i \pi_j}{\pi_{i,j} \pi_i \pi_j} y_i y_j \right]. \quad (4)$$

Hierbei gilt zu beachten, dass π_i bzw. $\pi_{i,j}$ in den Varianzen im Nenner vertreten ist. Umso geringer der Umfang der Stichprobe φ , umso kleiner wird auch π_i bzw. $\pi_{i,j}$, was entsprechend dazu führt, dass die Varianz größer wird. Für die Varianzen gibt es neben den hier dargestellten Schätzungen auch noch die exakten Berechnungen für die allerdings Informationen über die gesamte Grundgesamtheit Ω verfügbar sein muss (Vergleiche hierzu [Münnich et al., 2013, S. 154, 155]).

Der Vorteil des Horvitz-Thompson-Schätzers liegt darin, dass dieser nur auf zwei Eigenschaften des Stichprobenverfahrens beruht. Auf der einen Seite wird verlangt, dass kein Element doppelt in die Stichprobe gelangen kann und auf der anderen Seite, dass die Inklusionswahrscheinlichkeit π_i für jedes Element i aus der Grundgesamtheit Ω größer 0 ist ($\pi_i > 0, i \in \{1, 2, \dots, N\}$). Sind diese Annahmen erfüllt, so ist der Horvitz-Thompson-Schätzer Erwartungstreu. Diese wichtige Eigenschaft des Schätzers wird als Horvitz-Thompson-Theorem bezeichnet.

3.1.2 Verallgemeinerter Regressionsschätzer

Bei Stichprobenerhebungen liegen oft neben der zu untersuchenden Variable, noch weitere Informationen vor. Solche Informationen, auch Hilfsinformation genannt, können besonders gut eingebracht werden, wenn sie für die gesamte Grundgesamtheit verfügbar sind. Mögliche Hilfsinformationen sind beispielsweise Werte aus benachbarten Gebieten, Merkmalausprägungen aus der Vergangenheit oder Werte von Hilfsvariablen, die stark mit interessierenden Merkmal korreliert sind (Vergleiche [Dieterle, 2010, S. 1213]). So könnte in dem Beispiel mit der Schweinebestandsschätzung beispielsweise der Schweinebestand der letzten Totalerhebung oder der jährliche Umsatz des Betriebes als Hilfsinformation dienen.

Der verbreitetste Schätzer, der sich solche zusätzlichen Informationen zu Nutze machen kann ist der verallgemeinerte Regressionsschätzer (GREG =generalized regression estimator). Dieser verallgemeinerte Regressionsschätzer benutzt die Informationen eines Regressionsmodells zur Korrektur eines Horvitz-Thompson-Schätzwertes und fällt damit in die Kategorie der Modell unterstützenden Verfahren. Dieser verallgemeinerte Regressionsschätzer sieht für die Schätzung eines Totalwertes wie folgt aus.

$$\hat{\tau}_Y^{GREG} = \hat{\tau}_Y^{HT} + (\tau_X - \hat{\tau}_X^{HT})\hat{\beta}, \quad (5)$$

wobei $\hat{\beta}$ über

$$\hat{\beta} = \left(\sum_{i \in \varphi} \omega_i x_i' x_i \right)^{-1} \sum_{i \in \varphi} \omega_i x_i' y_i \quad (6)$$

geschätzt wird.

Der $\hat{\beta}$ -Vektor ist hierbei die Lösung des klassischen Kleinste-Quadrat-Schätzung. Dabei werden zusätzliche Designgewichte verwendet, was zu einer asymptotisch unverzerrten Schätzung des $\hat{\beta}$ -Vektors führt. Die β -Schätzung muss hierbei nicht notwendigerweise über eines linearen Regressionmodells durchgeführt werden, sondern kann auch durch ein gemischtes Modell [siehe Kapitel 5] erfolgen. x_i stellt die Hilfsinformationen des Zeilenvektors der i -ten Beobachtung dar. τ_x steht für die Totalwerte der Hilfinformation. Sind die klassischen Annahmen des Regressionsmodells erfüllt, so resultiert eine unverzerrte Schätzung für $\hat{\tau}_Y^{GREG}$.

Man sieht also, dass bei der Schätzung mit dem GREG-Verfahren der Fehler, der bei der alleinigen Schätzung von $\hat{\tau}_Y^{HT}$ entstehen kann, über den Term $(\tau_X - \hat{\tau}_X^{HT})\hat{\beta}$ korrigiert wird. Liegt die Schätzung für $\hat{\tau}_X^{HT}$ stark neben dem wahren Wert τ_X , so ist der Korrekturterm hoch und dem entsprechend wird $\hat{\beta}$ stark gewichtet. Liegt die Schätzung für τ_X genau richtig, so entspricht der GREG-Schätzer genau dem HT-Schätzer, da der Term $\tau_X - \hat{\tau}_X^{HT} = 0$ gilt

und damit der ganze Korrekturterm 0 wird.
Die Varianz des GREG-Schätzers lässt sich über

$$\hat{V}(\hat{\tau}_Y^{GREG}) = N^2 \cdot \frac{s^2}{n} \left(1 - \frac{n}{N}\right) \quad (7)$$

schätzen, wobei

$$s^2 = \frac{1}{n-1} \sum_{i \in \varphi} \varepsilon_i^2 \quad (8)$$

mit $\varepsilon_i = y_i - x_i \hat{\beta}$ gilt.

Auch hier ist, wie bei dem Horvitz-Thompson Schätzer, der Umfang der Stichprobe φ im Nenner vertreten, sodass die Varianz der Schätzung mit sinkendem Stichprobenumfang zunimmt.

4 Grundlagen der Small Area-Schätzung

4.1 Begriffliche Grundlagen

„Den Begriff Small Area [...] definiert Rao als inhaltlich oder geografisch abgegrenzte Subpopulation, für die eine direkte Schätzung“ [Dieterle, 2010, S. 1212], wie beispielsweise mit dem HT- oder GREG-Schätzer, aufgrund einer zu geringen Stichprobe nur mit hohen Standardfehlern möglich ist. Auch kann eine Schätzung gar nicht möglich sein, wenn die inhaltlich oder geografisch abgegrenzte Subpopulation überhaupt nicht in der Stichprobe enthalten ist. Im folgenden werden aus Gründen der Einfachheit Gebiete und Subgruppen unter dem Begriff *Area* vereint.

Ausgehend von der Grundgesamtheit Ω lässt sich diese in D Teilgesamtheiten aufteilen, so ließe sich beispielsweise das Bundesland Bayern in seine „71 Gemeinden“ [Bayrische Staatsregierung, 2016] aufteilen. Entsprechend gilt:

$$\Omega = \dot{\bigcup}_{d=1, \dots, D} \Omega_d \quad (9)$$

Ω_d bezeichnet die Teilgesamtheit, die aus der Area d stammt. Die verschiedenen Umfänge aus den disjunkten D Areas werden mit N_1, \dots, N_D bezeichnet. Entsprechend gilt $N = \sum_{i=1}^D N_i$.

Dabei muss die Stichprobe φ , genauso wie die Grundgesamtheit Ω , auf die D Areas aufgeteilt werden. Hierbei bezeichnet φ_d den Teil der Stichprobe, welcher aus der Area d stammt. So gilt

$$\varphi = \dot{\bigcup}_{d=1, \dots, D} \varphi_d. \quad (10)$$

n_d bezeichnet nun den Umfang der Stichprobe aus der Area d . Mathematisch lässt sich dies mit einer Indikatorvariable $z_{di} = I(i \in \varphi_d)$ ausdrücken. n_d lässt sich also als

$$n_d = \sum_{i \in \Omega} z_{di} \quad (11)$$

schreiben. Es gilt also, analog zu der Grundgesamtheit, $n = \sum_{i=1}^D n_i$. Da es vorkommen kann, dass Areas erst nach der Ziehung der Stichprobe definiert werden, ist es möglich, dass sich die Stichprobenumfänge in den verschiedenen Areas stark unterscheiden. Werden die Areas erst nach der Ziehung der Stichprobe definiert spricht man von *Unplanned Areas*. Das Gegenteil der Unplanned Areas bezeichnet man mit *Planned Areas*, bei denen die Stichprobenumfänge fest und vorher festgelegt sind. Hierbei sind Planned Areas den Unplanned Areas vorzuziehen. Dies liegt daran, dass bei Unplanned Areas der

Fall eintreten kann, dass bestimmte Areas überhaupt nicht in die Stichprobe gelangen und „in diesem Fall [...] designbasierte Verfahren nicht mehr verwendet werden“ [Münnich et al., 2013, S. 158] können. Letzteres traf auch im Falle der Schätzung von Schweinebeständen ein, da die Stichprobe auf Länderebene ausgelegt war, gab es Kreise, die überhaupt nicht in der Stichprobe enthalten waren.

Durch die Aufteilung der Grundgesamtheit in d Areas, werden auch die Total- und Mittelwerte für die verschiedenen Areas neu berechnet. Dies funktioniert komplett Analog zu der Berechnung der Grundgesamtheit, nur dass die Beobachtungen auf die zu interessierenden Areas beschränkt werden. So ist der areaspezifische Totalwert die Summe der Merkmalsausprägungen der Area d .

$$\tau_{Y,d} = \sum_{i \in \Omega_d} y_i \quad (12)$$

Dementsprechend ist der areaspezifische Mittelwert über

$$\mu_{Y,d} = \frac{1}{N_d} \sum_{i \in \Omega_d} y_i \quad (13)$$

definiert.

4.2 Designbasierte Area-Schätzung

Die bereits vorgestellten Schätzverfahren aus Kapitel drei lassen sich entsprechend für Schätzungen in den verschiedenen Areas erweitern. So ist beispielsweise die Schätzung des Total- bzw. Mittelwertes komplett analog zu den Formeln 1 und 2, nur dass hierbei die Beobachtungen entsprechend nicht aus der Stichprobe über die Grundgesamtheit φ stammen, sondern aus der Teilstichprobe φ_d . So wird der Mittelwert für eine bestimmte Area über

$$\hat{\mu}_{Y,d}^{HT} = \frac{1}{\hat{N}_d} \sum_{i \in \varphi_d} \frac{y_i}{\pi_i} = \frac{1}{\hat{N}_d} \sum_{i \in \varphi_d} y_i \omega_i \quad (14)$$

geschätzt, wobei sich N_d über $\sum_{i \in \varphi_d} \frac{1}{\pi_i}$ schätzen lässt. Analog dazu lässt sich der Areaspezifische Totalwert über

$$\hat{\tau}_{Y,d}^{HT} = \sum_{i \in \varphi_d} \frac{y_i}{\pi_i} = \sum_{i \in \varphi_d} y_i \omega_i \quad (15)$$

schätzen.

Auf unser Beispiel angewandt, wird die Anzahl der Schweine in einem Landkreis also auf Basis der Stichprobenerhebung aus diesem Landkreis geschätzt.

Hier fällt der große Nachteil von designbasierten Schätzverfahren auf. Ist ein Landkreis nicht in der Stichprobe enthalten, was schnell der Fall sein kann, wenn man überlegt, dass die Stichprobe eigentlich auf Länderebene ausgelegt war, so ist eine direkte Schätzung der Schweinebestände für diese Landkreise nicht möglich, so dass für diese Landkreise ein alternatives Schätzverfahren von Nöten ist.

Bei Regressionsschätzungen, wie dem GREG, gibt es zwei verschiedene Möglichkeiten, diese auf d Areas zu erweitern. Die zwei Möglichkeiten unterscheiden sich durch die Auswahl der Daten, welche in das Modell mitaufgenommen werden.

Separate Betrachtung Hierbei wird in jeder Area separat eine Regressionschätzung durchgeführt. Es werden also nur die Daten der Area d zur Schätzung von $\hat{\beta}$ genutzt. Hierbei kann natürlich auch wieder der Fall eintreten, dass die Schätzungen aufgrund sehr kleiner Teilstichprobenumfänge bzw. nicht vorhandener Teilstichprobenumfänge sehr ungenau werden bzw. gar nicht möglich sind.

Indirekte Schätzung Bei dieser zweiten Variante wird die Schätzung des Modells auf Basis aller verfügbaren Informationen der Stichprobe über die verschiedenen D Areas hinweg durchgeführt. Hierbei wird β also unter Verwendung aller Informationen der Stichprobe geschätzt und beschränkt sich nicht nur auf die betrachtete Area d. Das Modell wird also „auf Basis der Informationen von übergeordnete[n] Areas gebildet“ [Münnich et al., 2013, S. 159].

Bei dem deutschen Zensus gab es die Diskussion, ob die Regressionkoeffizienten auf Basis der Grundgesamtheit, also Deutschland, oder auf Basis von Gemeinden, Kreisen oder Bundesländern, also Teilgesamtheiten, erfolgen sollte. Es wurde sich für eine direkte Schätzung entschieden, sodass die Regressionskoeffizienten auf Informationsbasis der jeweiligen Gemeinden, und nicht etwa auf der Basis von ganz Deutschland, geschätzt werden.

4.3 Synthetische Schätzungen

Der große Nachteil, bei designbasierten Verfahren wie dem HT und modellunterstützten Verfahren wie dem GREG, liegt darin, dass die Varianz der Schätzfunktion unvertretbar groß wird, wenn die Teilstichprobenumfänge zu

klein werden. Um dieses Problem zu umgehen ist es möglich anstatt der direkten Schätzung eine indirekte, modellbasierte Schätzung zu verwenden. So lässt sich beispielsweise ein synthetischer Schätzer anwenden. Diese Schätzer nutzen Informationen aus einer übergeordneten Einheit. Dabei wird davon ausgegangen, dass die untergeordneten Areas dieselbe Charakteristik wie die übergeordnete Einheit aufweisen. Informationen aus der übergeordneten Gesamtheit werden also synthetisch auf die einzelnen Areas heruntergebrochen. Es ist also fundamental wichtig, dass die Charakteristik in der Grundgesamtheit die gleiche ist wie in den untergeordneten Areas. Ist dies nicht der Fall ist die Modellannahme, die bei synthetischen Schätzern getroffen wird, verletzt und man spricht man von einer Fehlspezifikation des Modells. Dementsprechend wären die Schätzungen des synthetischen Schätzers nicht valide.

Der *National Sample Mean* (NSM) wird hierbei gerne als klassischer Benchmark herangezogen. Dieser berechnet sich über

$$\hat{\mu}_{Y,d}^{NSM} = \frac{\sum_{i \in \varphi_d} w_i y_i}{\sum_{i \in \varphi_d} w_i} \quad d = 1, \dots, D. \quad (16)$$

Da man davon ausgeht, dass die Areas sich untereinander nicht zu stark unterscheiden, müssten sich die Mittelwerte der einzelnen Areas ähnlich sein und es müsste entsprechend gelten

$$\hat{\mu}_{Y,d}^{NSM} = \hat{\mu}_Y \quad d = 1, \dots, D. \quad (17)$$

Ist dies nicht der Fall, so scheint die grundlegende Annahme, dass sich die Charakteristik in der Grundgesamtheit die gleiche ist wie in den untergeordneten Areas, verletzt. Dementsprechend wären die Schätzungen nicht mehr valide.

Alternativ lässt sich auch ein ungewichteter NSM berechnen, falls die Umfänge der Teilgesamtheiten unbekannt sein sollten. Auch lässt sich der NSM grundsätzlich auch auf Totalwerte umrechnen, wenn die Umfänge der verschiedenen Areas bzw. Schätzungen dafür vorliegen.

Diese synthetische Vorgehensweise lässt sich in viele Richtungen verallgemeinern. Wird nun von einem Regressionsmodell, welches sich auf alle Areas übertragen lässt, ausgegangen, so erhält man als regressionssynthetischen Schätzer

$$\hat{\tau}_{Y,d}^{SYNR} = \tau_{X,d} \hat{\beta}. \quad (18)$$

Hierbei lässt sich $\hat{\beta}$ über den gewichteten KQ-Schätzer (Vergleiche Formel 6) berechnen, dabei gehen alle Informationen der Stichprobe in die β -Schätzung mit ein. Dieser synthetische Multilevel-Regressionschätzer wird im folgenden

mit SYNM bezeichnet werden. Auch hier ist, wie bei dem GREG-Schätzer, ein gemischtes Modell [siehe Kapitel 5] zur Schätzung von $\hat{\beta}$ möglich. Es sind aber auch komplett andere Modelle möglich, wie beispielsweise einem Nicht-Linearen Modell oder einem Verhältniss-Modell. Letzteres fand bei der Schätzung von Schweinebeständen seine Anwendung. So wurde anstatt des $\hat{\beta}$ eine Variable a_{k2003} eingeführt, welche den „Anteil der Schweine in einem Kreis an allen Schweinen im jeweiligen Bundesland im Jahr 2003“ [Dieterle, 2010] darstellt. Auf diesen speziellen sythetischen Schätzer soll in Kapitel 7 noch genauer eingegangen werden.

4.4 Zusammengesetzte Schätzer

Es haben sowohl sythetische, als auch designbasierte Schätzverfahren verschiedene Vor- und Nachteile. So ist bei synthetischen Schätzverfahren die Gefahr einer Verzerrung relativ groß und zwar genau dann, wenn die Modellspezifikation nicht mehr zutreffend ist. Dafür ist die Schätzvarianz bei den synthetischen Methoden sehr gering. Bei designbasierten Ansätzen sieht dies ganz anders aus, so sind diese Mehtoden im allgemeinen (asymptotisch) designunverzerrt. Bei kleinen (Teil-) Stichprobenumfängen kann es allerdings vorkommen, dass die Designvarianzen unakzeptabel hoch sind.

Man sieht also, dass jedes Verfahren aus den zwei Kategorien einen klaren Vorteil, aber auch einen klaren Nachteil zu bieten hat. Diese sind in Tabelle 1 nocheinmal zusammenfassend dargestellt.

	Designbasierte Methoden	Synthetische Methoden
<i>Bias</i>	Design unverzerrt	Modellbedingter Bias
<i>Design Varianz</i>	Hoch, falls n_d klein	Gering, n_d mit schwachem Einfluss
<i>MSE</i>	$MSE \approx$ Varianz	Modell-Bias eventuell sehr dominant

Tabelle 1: Vor- und Nachteile der zwei Schätzmethoden in Anlehnung an [Münnich et al., 2013, S. 162]

Es ist also ersichtlich, dass die Stärke des einen Schätzverfahrens die Schwäche des anderen Verfahrens ist. Möchte man nun diese Vorteile maximieren bzw. die Nachteile minimieren, so lässt sich dies über eine Linearkombination der beiden Komponenten erreichen. Dies führt zu einem sogenannten *zusammengesetzten Schätzer*, dessen Sinn darin besteht die Vorteile der jeweiligen Verfahren zu nutzen und deren Nachteile zu reduzieren. Dieser

zusammengesetzte Schätzer hat die folgende Form

$$\hat{\tau}_{Y,d}^{COMP} = \delta_{Y,d} \hat{\tau}_d^{DES} + (1 - \delta_d) \hat{\tau}_{Y,d}^{SYN}, \quad (19)$$

wobei $\delta_d \in [0; 1]$ gilt. Dabei steht $\hat{\tau}_{Y,d}^{DES}$ für den geschätzten Totalwert, der über eine designbasierte Methode geschätzt wurde und $\hat{\tau}_{Y,d}^{SYN}$ für den synthetisch geschätzten Totalwert. Allerdings weist auch der zusammengesetzte Schätzer Probleme bei der Wahl der Gewichtungsfaktoren δ_d auf. Hierbei gibt es grundsätzlich zwei Verfahren zur Wahl der Gewichte.

Fallzahlabhängige Gewichtung Bei der Entscheidung für eine Fallzahlabhängige Gewichtung, werden einfache Kennwerte, wie beispielsweise der geschätzte Populationsumfang \hat{N} oder der geschätzte Totalwert einer Hilfsvariable $\hat{\tau}_X$ herangezogen, um so die Gewichte geeignet wählen zu können. So wählt [Drew et al., 1982] beispielsweise:

$$\delta_d^{DSC} = \begin{cases} 1, & \hat{N}_d \geq \alpha N_d \\ \frac{\hat{N}_d}{\alpha N_d}, & \text{sonst} \end{cases}$$

Hierbei lassen sich mögliche Unterrepräsentationen der Stichprobe, welche sich in der Unterschätzung von N_d ausdrücken, aus Area d kompensieren. Ist die Schätzung für N_d um ein α -faches höher als das wahre N_d , so entspricht der zusammengesetzte Schätzer genau dem direkten Schätzer.

Auch hier besteht das Problem, dass die Konstante $\alpha > 0$ erstmalig geeignet gewählt werden muss. Auch werden bei Methoden wie diesen weder die Varianz- noch die MSE-Größen der zwei Komponenten berücksichtigt.

Optimale Gewichtung Bei der optimalen Gewichtung, wird der MSE des zusammengesetzten Schätzers betrachtet, welcher, bei Vernachlässigung der Kovarianz, die Form

$$MSE(\tau_d^{COMP}) = \delta_d^2 MSE(\hat{\tau}_d^{DIR}) + (1 - \delta_d)^2 MSE(\hat{\tau}_d^{SYN}) \quad (20)$$

hat. So lässt sich durch die Minimierung des mittlerern quadratischen Fehlers das optimale Gewicht δ_k^* bestimmen.

$$\delta_k^* = \frac{MSE(\hat{X}_{k2007}^{SYN})}{MSE(\hat{X}_{k2007}^{HT}) + MSE(\hat{X}_{k2007}^{SYN})} \quad (21)$$

Umso kleiner der mittlere quadratische Fehler des synthetischen Schätzers im Vergleich zu dem mittlere quadratische Fehler des direkten Schätzers, umso

stärker wird der synthetische Schätzer gewichtet und geht dementsprechend auch stärker in die Schätzung mit ein. Bei einem optimalen Gewicht δ_k^* von 0 oder 1 entspricht der zusammengesetzte Schätzer genau dem präziseren Schätzer von beiden.

5 Spezielle Small Area-Schätzmethoden

In diesem Abschnitt, der sich an [Münnich et al., 2013, S. 163 - 166] anlehnt, sollen zwei zusammengesetzte Schätzer, welche die Eigenschaft der MSE-Minimalität erfüllen, vorgestellt werden. Hierbei handelt es sich um spezielle zusammengesetzte Schätzer, welche als Spezialfall eines allgemeinen gemischten Modells gesehen werden. Im folgenden werden die grundlegenden Annahmen, welche man für diese gemischten Modelle trifft erläutert.

Ein gemischtes Modell hat die Form

$$y = x\beta + zv + \epsilon \quad (22)$$

Dabei sind e , x , y wie beim klassischen Regressionsmodell definiert sind, z ist eine bekannte $n \times k$ Matrix und v stellt einen $k \times 1$ Vektor der zufälligen Effekte dar. Im Small Area Kontext wird z häufig als Indikator-Matrix, welche die Beobachtungen den entsprechenden Areas zuordnet, gewählt. Desweiteren wird angenommen, dass ϵ und v stochastisch unabhängig sind und sowohl ϵ , als auch v multivariat normalverteilt sind mit μ_v bzw. $\mu_\epsilon = 0$, sowie Varianz-Kovarianz-Matrizen Σ_v und Σ_ϵ . Bei gemischten Modellen wird, anders als bei linearen Regressionsmodellen, angenommen, dass die Variation der interessierenden Variable nicht nur durch die festen β Effekte erklärt wird, sondern auch durch zufällige Effekte. Dabei werden die zufälligen Effekte meist auf einer höheren Aggregationsebene, als den beobachteten Einheiten angesetzt. Auf das Beispiel mit der Schweinsbestandsschätzung übertragen, bedeutet dies, dass sich Schweinehaltende Betriebe innerhalb eines Landkreises ähnlicher sind als zwischen unterschiedlichen Bundesländern. Unter der Gültigkeit der Annahmen des gemischten Modells, lässt sich β über

$$\hat{\beta} = (x'v^{-1}x)^{-1} x'v^{-1}y \quad (23)$$

schätzen, wobei

$$v = \Sigma_\epsilon + z\Sigma_v z' \quad (24)$$

gilt. Im Anwendungsgebiet der Small Area-Schätzung, wird zudem noch „angenommen, dass ϵ sowie v identisch verteilt und stochastisch unabhängig sind. Damit reduzieren sich Σ_ϵ und Σ_v zu $\sigma_\epsilon I$ und $\sigma_v I$, wobei I die Einheitsmatrix [...] darstellt“ [Münnich et al., 2013, S. 164].

5.1 Battese-Harter-Fuller-Schätzer

Der Battese-Harter-Fuller-Schätzer (BHF) nach [Battese et al., 1988] dient zur Schätzung von Mittelwerten interessierender Variablen auf Ebene der Small Areas. Hierbei verwendet dieser Schätzer Informationen auf kleinster disaggregierter (Individual-)ebene, weshalb von einem Unit-Level-Modell gesprochen wird. Die Schätzung der Area Mittelwerte mit dem BHF hat die Form

$$\hat{\mu}_{Y,d}^{BHF} = \delta_d \left(\hat{\mu}_{Y,d}^{HT} + (\bar{X}_d - \mu_{X,d}^{HT}) \hat{\beta} \right) + (1 - \delta_d) \bar{X}_d \hat{\beta}, \quad (25)$$

wobei

$$\delta_d = \frac{\sigma_v^2}{\sigma_v^2 + \frac{\sigma_\epsilon^2}{n_d}} \quad (26)$$

gilt. \bar{X}_d bezeichnet dabei die Durchschnittswerte der Hilfsmerkmale in jeder Area d , wobei \bar{X}_d schließlich $\mu_{X,d}$ entspricht. $\mu_{X,d}^{HT}$ bezeichnet dabei die geschätzten Durchschnittswerte der Hilfsmerkmale aus jeder Area.

Die Zusammensetzung aus einem direktem, sowie einem synthetischen Schätzer wird in dieser Darstellungsform leicht ersichtlich. Dabei handelt es sich bei dem ersten Teil, dem direkten Schätzer, um einen Multi-Level GREG Schätzer und bei dem synthetischen Schätzer um einen SYNM Schätzer. Der BHF-Schätzer „erfüllt die Eigenschaft, bester unverzerrter Prädiktor unter den linearen Schätzern zu sein“ [Münnich et al., 2013]. Bei der praktischen Anwendung dieses Schätzverfahrens sind die Varianzkomponenten σ_v und σ_ϵ meist unbekannt, sodass diese Komponenten wieder geschätzt werden müssen. Diese Schätzungen sind in den heute gängigen Software-Programmen bereits in die Schätzung des gemischten Modells integriert. Werden nun anstatt der wahren σ_v und σ_ϵ die geschätzten $\hat{\sigma}_v$ und $\hat{\sigma}_\epsilon$ verwendet, wird auch δ_d zu $\hat{\delta}_d$. Werden die geschätzten Varianzen verwendet, so ist der BHF-Schätzer der „empirisch bester unverzerrter Prädiktor unter den linearen Schätzern“ [Münnich et al., 2013, S. 165].

5.2 Fay-Herriot Schätzer

Auch der Fay-Herriot Schätzer (FH) nach [Fay and Herriot, 1979] dient zur Schätzung von Mittelwerten interessierender Variablen auf Ebene der Small Areas. Der Unterschied zwischen dem FH- und dem BHF-Schätzer liegt in der Art, wie die zwei Schätzer die Informationen verwenden. Wie bereits in vorherigem Abschnitt festgestellt, verwendet der BHF-Schätzer Informationen auf Individualebene. Ganz anders funktioniert hier der FH-Schätzer, da dieser Informationen nur auf aggregierter Ebene verwendet, weshalb auch

von einem Area-Level-Modell gesprochen wird. Der Vorteil eines Area-Level-Modells gegenüber einem Unit-Level-Modell liegt darin, dass es für die Area-Level-Modelle genügt die aggregierten Daten vorliegen zu haben und diese Modelle entsprechend mit weniger Information auskommen. Dies ist besonders im Hinblick auf den Datenschutz vorteilhaft, da nur aggregierte und keine Mikrodaten von Nöten sind. Zudem lassen sich bei Area-Modellen Daten aus mehreren Quellen leichter zusammenführen und müssen nicht erst aufwendig zusammengeführt werden, wie zum Beispiel mit dem Verfahren des Record Linkage. Allerdings führt das aber auch dazu, dass der FH-Schätzer „in vielen Fällen schlechtere Schätzungen liefert als der BHF“ [Münnich et al., 2013, S. 166].

Die Schätzung der Mittelwerten mit dem FH-Schätzer lässt sich über

$$\hat{\mu}_{Y,d}^{FH} = \delta_d \mu_{Y,d}^{HT} + (1 - \delta_d) \bar{X}_d \hat{\beta} \quad (27)$$

darstellen, wobei sich δ_d genauso wie in Formel 26 berechnet. Auch hier lässt sich die zusammengesetzte Form, von direktem und synthetischem Schätzer gut erkennen. Hierbei ist der direkte Schätzer ein HT-Schätzer und der synthetische ein SYNM-Schätzer. Auch der FH-Schätzer ist ein empirisch bester unverzerrter Prädiktor unter den linearen Schätzern.

6 Vergleich der verschiedenen Schätzverfahren nach [Dieterle, 2010]

In diesem Abschnitt der Arbeit, der sich an [Dieterle, 2010, S. 1214-1218] anlehnt, sollen die verschiedenen Schätzverfahren, anhand des Beispiels der Schweinebestandsschätzung, miteinander verglichen werden. Dabei werden die verschiedenen Verfahren zunächst noch einmal auf Schweinebestandsschätzung bezogen vorgestellt und anschließend miteinander verglichen.

6.1 Anwendung der Schätzverfahren im Beispiel der Schweinebestandsschätzung

Direktes Schätzverfahren Im Mai 2007 erreichten etwas mehr als 80'000 Betriebe in 420 Landkreisen die Erfassungsgrenze der Agrarstrukturerhebung, wovon dann knapp 21'000 Betriebe aus 399 Landkreisen in der Stichprobenmaterial enthalten waren. Für die 21 Landkreise, die nicht in der Stichprobe enthalten sind, ist eine direkte Schätzung der Schweinebestände folglich nicht möglich. Als direktes Schätzverfahren dient das HT-Schätzverfahren [Vergleiche Formel 1], sodass sich die Anzahl der Schweine in einem bestimmten Landkreis über

$$\hat{\tau}_{Y,d2007}^{HT} = \sum_{i \in \varphi_d} \frac{y_i}{\pi_i} \quad (28)$$

schätzen lässt, wobei y_i ($i = 1, \dots, n$) dem Schweinebestand der Stichprobe i aus dem Landkreis d ($d = 1, \dots, 399$) entspricht. Die Schätzung berücksichtigte ursprünglich noch 432 verschiedene Schichten, unter anderem zum Beispiel Größenklassen und Hauptproduktionsrichtung. Diese wurden in diesem Teil der Arbeit weggelassen, werden aber in [Dieterle, 2010, S. 1214] ausführlich beschrieben

Syntehtischer Schätzer Bei diesem Verfahren, welches sich Informationen aus einer übergeordneten Einheit zu Nutze macht, wird der Anteil der Schweine in den einzelnen Kreisen am gesamten Bundesland aus dem Jahr 2003 als Hilfsinformation genutzt. Dabei wurden die Ergebnisse der Kreise, deren Gebiete sich von 2003 auf 2007 geändert haben, mithilfe eines flächenbezogenen Umrechnungsfaktors auf die Kreisgrenzen des Jahres 2007 umgerechnet, sodass die Schweinebestände aus dem Jahr 2003 an die Gebiete des Jahres 2007 angepasst wurden. Dabei lässt sich der Anteil der Schweine

in den einzelnen Kreisen am gesamten Schweinebestand des Landes aus dem Jahr 2003 ausdrücken als:

$$a_{k2003} = \frac{X_{k2003}}{X_{l2003}}, \quad (29)$$

wobei X_{k2003} für die Anzahl der Schweine in dem Kreis in 2003 steht und X_{l2003} für die Anzahl der Schweine in dem jeweiligen Bundesland im Jahre 2003. Als Hilfsinformation dient die über das HT-Schätzverfahren geschätzte Anzahl der Schweine im Bundesland l $\hat{\tau}_{l2007}^{HT}$. So erhält man als synthetischen Schätzer für die Schweinebestände auf Kreisebene:

$$\hat{\tau}_{Y,d2007}^{SYN} = a_{k2003} \hat{\tau}_{X,d2007}^{HT}. \quad (30)$$

Es wird also davon ausgegangen, dass sich die Verhältnisse a_{k2003} und a_{k2007} nicht sonderlich unterscheiden. Macht eine Kreis also 10 % des Schweinebestandes in einem Bundesland im Jahr 2003 aus, so wird davon ausgegangen, dass dieser Kreis im Jahr 2007 immer noch etwa 10 % des Schweinebestandes des Bundeslandes ausmacht.

Zusammengesetzter Schätzer Dieser Schätzer setzt sich aus den zwei oben vorgestellten Verfahren zusammen. Also einmal dem direkten HT-Schätzer und einmal dem synthetischen Schätzer. Folglich hat er folgende Form:

$$\hat{\tau}_{Y,d2007}^Z = \delta_d \cdot \hat{\tau}_{Y,d2007}^{HT} + (1 - \delta_d) \cdot \hat{\tau}_{Y,d2007}^{SYN} \quad (31)$$

Dabei legt δ_d den Anteil des direkten Schätzers fest und $(1 - \delta_d)$ entsprechend den Anteil des synthetischen Schätzers. Hierbei wird δ_d durch die Minimierung des mittleren quadratischen Fehler bestimmt (Vergleiche Kapitel 4).

6.2 Verfahren zum Vergleich der verschiedenen Schätzverfahren nach [Dieterle, 2010]

Zum Vergleich der Genauigkeit der verschiedenen Verfahren werden die relativen Standardfehler bzw. die relative Wurzel aus dem mittleren quadratischen Fehler benötigt. Für den direkten Schätzer wird der relative Standardfehler benötigt, der angibt wie stark der Standardfehler der Schätzung ist, dieser berechnet sich wie folgt:

$$RSTF(HT) = \frac{S_{\hat{\tau}_{d2007}^{HT}}}{\hat{\tau}_{d2007}^{HT}} \quad (32)$$

Dabei steht $\hat{\tau}_{d2007}^{HT}$ für den geschätzten Schweinebestand des Landkreises d und $s_{\hat{\tau}_{d2007}^{HT}}$ für die Standardabweichung der Schätzung. Hierbei gilt, umso kleiner der relative Standardfehler, umso besser die Schätzung des direkten Schätzers.

Bei dem synthetischen und zusammengesetzten Schätzer ist eine Bewertung der Genauigkeit nicht alleine anhand des relativen Standardfehler sinnvoll, da bei diesen Verfahren, durch die Hinzunahme von Hilfsinformationen eine Verzerrung entstehen kann, welche bei dem relativen Standardfehler nicht berücksichtigt wird. Hierbei entsteht die Gefahr der Verzerrung durch eine Veränderung der Anteile zwischen den Jahren 2003 (a_{k2003}) und 2007 (a_{k2007}). So weißt der synthetische Schätzer beispielsweise für Kreise, die 2003 Schweine hielten, aber 2007 nicht mehr, trotzdem noch Schweine für 2007 auf (Vergleiche Tabelle 3 Landkreis H). „Die Genauigkeit eines verzerrten Schätzers wird durch den mittleren quadratischen Fehler (mean squared error, MSE) gemessen, der sich aus der Varianz (stichprobenbedingter Fehler) und der Verzerrung (systematischer Fehler) zusammensetzt“ [Dieterle, 2010, S. 1216]. Analog zum relativen Standardfehler, existiert für verzerrte Schätzer die relative Wurzel aus dem mittleren quadratischen Fehler, der sich für den synthetischen (SYN) und zusammengesetzten (Z) Schätzer folgt berechnet:

$$RRNSE(SYN/Z) = \frac{\sqrt{MSE_{\hat{X}_{k2007}^{SYN/Z}}}}{\hat{X}_{k2007}^{SYN/Z}}, \quad (33)$$

wobei SYN/Z für entweder den synthetischen oder den zusammengesetzten Schätzer steht. Die Berechnung für beide Schätzverfahren die gleiche ist. Auch hier gilt, umso kleiner die relative Wurzel aus dem mittleren quadratischen Fehler, umso besser ist die Schätzung. In dem Bericht von [Dieterle, 2010] wird ein Ergebniss als gut bezeichnet, wenn die Wurzel aus dem mittleren quadratischen Fehler bzw. der relative Standardfehler unter 15% liegt.

6.3 Ergebnisse der Vergleiche

Zu dem Vergleich der drei Schätzverfahren sollen nun acht exemplarische Landkreise dienen. Die Ergebnisse der verschiedenen Schätzverfahren werden tabellarisch auf den nächsten Seiten dargestellt. Hierbei gilt zu beachten, dass der relative Standardfehler des direkten Schätzers, sowie die Wurzel des mittleren quadratischen Fehlers für synthetischen und zusammengesetzten Schätzer, unter dem Begriff *Relativer Standardfehler* zusammengefasst werden. In den folgenden Tabellen steht τ für den (wahren) Wert der Totaler-

hebung, $\hat{\tau}^{HT}$ für den geschätzten Totalwert des direkten Schätzers, $\hat{\tau}^{SYN}$ für den synthetisch geschätzten Totalwert und $\hat{\tau}^Z$ für den geschätzten Totalwert des zusammengesetzten Schätzers.

Zu Beginn werden die Ergebnisse des HT-Schätzverfahrens in Tabelle 2 dargestellt. Hierbei sind die Ergebnisse nach Höhe der relativen Standardfehler sortiert. Neben den bereits beschriebenen Elementen enthält Abbildung 2 noch die Totale Anzahl der Betriebe in dem Landkreis, sowie die Anzahl der Betriebe, die in der Stichprobe enthalten waren. Es lässt sich erkennen, dass umso kleiner der relative Standardfehler ist, umso kleiner auch die (relative) Abweichung des wahren Wertes vom geschätzten Wert ist. So ist die direkte Schätzung in den Landkreisen A und B sehr gut und die direkte Schätzung für die Landkreise E und F nicht besonders verlässlich. Zu Landkreis G, lässt sich keine direkte Schätzung ermitteln, da dieser Landkreis nicht in der Stichprobe enthalten war. Landkreis H ist ein Kreis, der keine Schweinehaltenden Betrieb hat, sodass dort auch keine direkte Schätzung möglich ist.

Landkreis	Schweinebestand 2007		Relativer Standardfehler $\hat{\tau}^{HT}$	Betriebe in der	
	τ	$\hat{\tau}^{HT}$		Totalerhebung	Stichprobe
A	94 414	94 680	0,2	81	47
B	830 303	850 517	5,1	1 530	382
C	7 416	6 190	10,0	11	2
D	12 153	15 522	22,0	91	21
E	4 216	4 864	36,4	72	17
F	2 457	3 147	50,1	47	10
G	714	.	.	4	0
H	0	.	.	0	0

Tabelle 2: Ausgewählte Ergebnisse des direkten Schätzers des Schweinebestandes 2007 auf Kreisebene in Anlehnung an [Dieterle, 2010, S. 1214]

Insgesamt schwankt der relative Standardfehler zwischen 0 und 97 %, wobei der Mittelwert bei 20 % liegt. Die Betrachtung der Verteilung der relativen Standardfehler, die grafisch durch Abbildung 1 unterstützt wird, erkennt man, dass der relative Standardfehler für 116 Landkreise unter 5 % liegt, in 170 Landkreisen unter 10 % liegt und insgesamt 229 Landkreise ($\hat{=}$ 55 %) unter 15 % liegen und damit brauchbare bzw. veröffentlichbare Ergebnisse liefert. Allerdings bedeutet das auch, dass 45 % der Ergebnisse sehr unsichere und damit nicht veröffentlichbare Ergebnisse liefert.

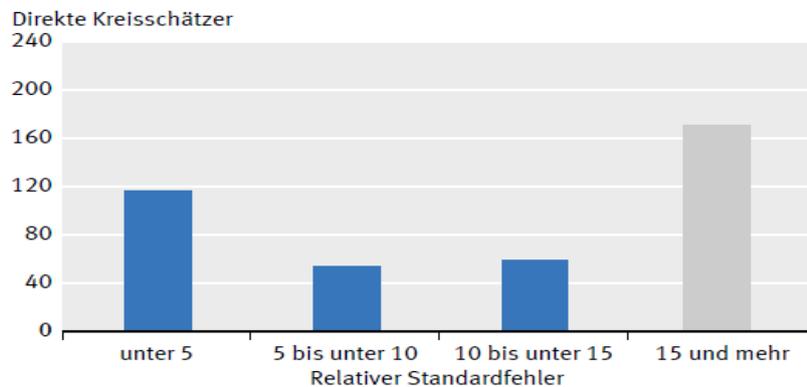


Abbildung 1: Verteilung des relativen relativen Standardfehlers des direkten Schätzers aus [Dieterle, 2010, S. 1215]

Tabelle 2 wird nun im nächsten Schritt um den synthetischen Schätzer, sowie dessen relative Wurzel aus dem mittleren quadratischen Fehler, erweitert und ist in Tabelle 3 dargestellt. Es lässt sich erkennen, dass der Synthetische Schätzer teils eine bessere Schätzung, zum Beispiel Landkreis C und D, und teils eine schlechtere Schätzung, zum Beispiel Landkreis A und B, als der direkte Schätzer liefert. Insgesamt scheint der synthetische Schätzer dann überlegen zu sein, wenn der direkte Schätzer hohe relative Standardfehler aufweist (Landkreis D und F). Auch ist mit dem synthetischen Schätzer eine Schätzung für den Landkreis J, der nicht in der Stichprobe enthalten war, möglich. Allerdings wird bei der Schätzung mit dem synthetischen Schätzer auch für Landkreis H, der im Jahr 2003 Schweinehaltende Betriebe hatte, aber im Jahr 2007 anscheinend nicht mehr, eine Schätzung abgegeben.

Landkreis	Schweinebestand 2007			Relativer Standardfehler	
	τ	$\hat{\tau}^{HT}$	$\hat{\tau}^{SYN}$	$\hat{\tau}^{HT}$	$\hat{\tau}^{SYN}$
A	94'414	94'680	95'669	0,2	1,4
B	830'303	850'517	803'528	5,1	3,7
C	7'416	6'190	8'260	10,0	11,4
D	12'153	15'522	13'790	22,0	14,6
E	4'216	4'864	5'565	36,4	24,2
F	2'457	3'147	2'605	50,1	8,4
G	714	.	1'269	.	43,2
H	0	.	2	.	100,2

Tabelle 3: Ausgewählte Ergebnisse des direkten und synthetischen Schätzers des Schweinebestandes 2007 auf Kreisebene in Anlehnung an [Dieterle, 2010, S. 1216]

Insgesamt schwankt die relative Wurzel aus dem mittleren quadratischen Fehler bei dem synthetischen Schätzer zwischen 0 und 191 %, wobei der Mittelwert bei 18 % und damit niedriger als bei dem direkten Schätzer (20 %) liegt. Betrachtet man die Abbildung 2, welche die Verteilung der relativen Wurzel der mittleren quadratischen Fehler für synthetischen Schätzer darstellt, so erkennt man, dass der Anteil der Schätzungen über 15 % im Vergleich zu dem direkten Schätzer stark zurück gegangen ist. Entsprechend steigt auch die Anzahl der veröffentlichbaren Ergebnisse von 55 % auf 69 %. Die relative Wurzel des mittleren quadratischen Fehlers liegt für 288 Landreise unter 15 %, für 227 Landkreise unter 10 % und für 141 Landkreise sogar unter 5 %. „Dies stellt im Vergleich zu dem direkten Schätzer eine leichte Verdichtung der besseren Ergebnisse dar“ [Dieterle, 2010, S. 1216].

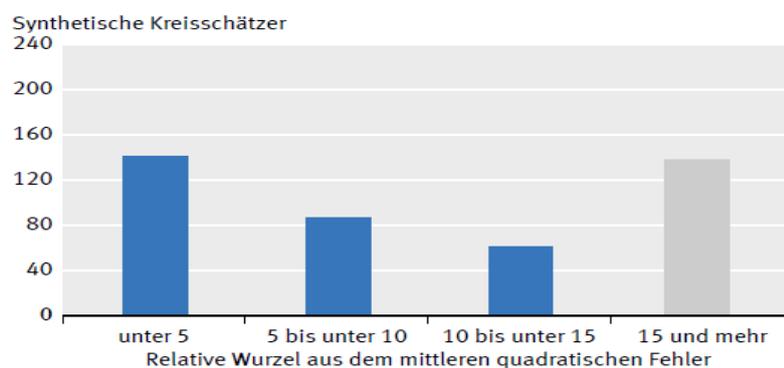


Abbildung 2: Verteilung der relativen Wurzel aus dem mittleren quadratischen Fehler des synthetischen Schätzers aus [Dieterle, 2010]

Im letzten Schritt soll nun Tabelle 3 um die Ergebnisse des zusammengesetzten Schätzers, sowie dessen relative Wurzel aus dem mittleren quadratischen Fehler und des Gewichtes δ für den direkten Schätzer, erweitert werden. Dargestellt in Abbildung 4. Für Landkreise, die nicht in der Stichprobe enthalten waren (J und K) entspricht der zusammengesetzte Schätzer genau dem synthetischen Schätzer ($\delta = 0$). Für den Landkreis A, dessen relativer Standardfehler sehr gering ist, besteht der zusammengesetzte Schätzer zu einem Großteil aus dem direkten Schätzer ($\delta = 0,98$). Der zusammengesetzte Schätzer bringt in einigen der beispielhaften Fälle eine deutliche Verbesserung der Schätzungen (zum Beispiel Landkreise B und C). Liegt das Gewicht δ nahe bei 0 oder 1, so führt der zusammengesetzte Schätzer zu keiner signifikanten Verbesserung des Ergebnisses (zum Beispiel Landkreise A und F).

Landkreis	Schweinebestand 2007				Relativer Standardfehler			Gewicht δ
	τ	$\hat{\tau}^{HT}$	$\hat{\tau}^{SYN}$	$\hat{\tau}^Z$	$\hat{\tau}^{HT}$	$\hat{\tau}^{SYN}$	$\hat{\tau}^Z$	
A	94'414	94'680	95'669	94'702	0,2	1,4	0,2	0,98
B	830'303	850'517	803'528	818'415	5,1	3,7	3,0	0,32
C	7'416	6'190	8'260	6'816	10,0	11,4	7,6	0,7
D	12'153	15'522	13'790	14'236	22,0	14,6	12,2	0,26
E	4'216	4'864	5'565	5'309	36,4	24,2	20,2	0,37
F	2'457	3'147	2'605	2'615	50,1	8,4	8,3	0,02
G	714	.	1 269	1269	.	43,2	43,2	0,00
H	0	.	2	2	.	100,2	100,2	0,00

Tabelle 4: Ausgewählte Ergebnisse der drei Schätzverfahren für den Schweinebestand 2007 auf Kreisebene in Anlehnung an [Dieterle, 2010, S. 1217]

Die Wurzel des mittleren quadratischen Fehlers schwankt zwischen 0 und rund 184 %, allerdings liegt der Durchschnitt mit 12 % weit unter dem Durchschnitt von synthetischem (18 %) und direktem Schätzer (20 %). Abbildung 3 stellt die Verteilung der relativen Wurzel der mittleren quadratischen Fehler für zusammengesetzten Schätzer dar. Der größte Teil der Schätzungen mit diesem Schätzer, nämlich 224 Landkreise ($\hat{=}$ 54%), liegt die relative Wurzel aus dem mittleren quadratischen Fehler unter 5 %. Die Menge der veröffentlichbaren Ergebnisse - gemessen an der 15 % Grenze - liegt für den direkten Schätzer bei ca. 80 %, was 337 Landkreisen entspricht. Durch die Verwendung des zusammengesetzten Schätzers wurde „die Anzahl veröffentlichungswürdiger Werte signifikant gesteigert“ [Dieterle, 2010, S. 1218].

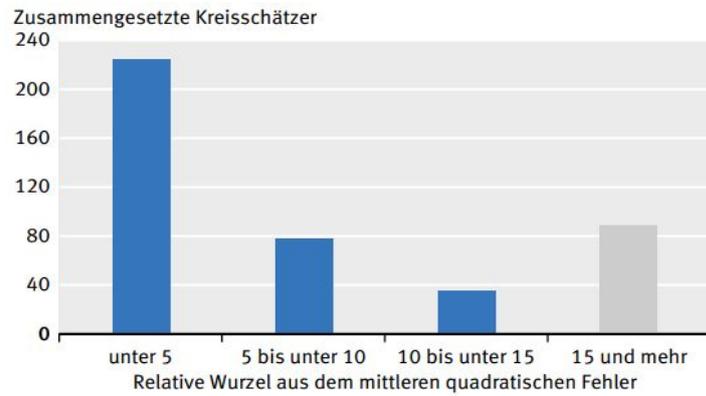


Abbildung 3: Verteilung der relativen Wurzel aus dem mittleren quadratischen Fehler des zusammengesetzten Schätzers aus [Dieterle, 2010, S. 1218]

7 Fazit und Ausblick

Zusammenfassend lässt sich sagen, dass bei dem Vergleich der drei Schätzverfahren, im Beispiel der Schweinebestandsschätzung, der direkte Schätzer wohl die schlechtesten Ergebnisse - gemessen an dem relativen Standardfehler bzw. der relativen Wurzel des mittleren quadratischen Fehlers - liefert und der zusammengesetzte Schätzer die besten Ergebnisse. Somit scheinen die Methoden der Small Area-Schätzung im Vergleich zu klassischen Schätzverfahren zumindest in diesem Fall wesentlich bessere Ergebnisse zu erzielen. Auch wenn der Anteil der unbefriedigenden Ergebnisse bei dem zusammengesetzten Schätzer am niedrigsten ist, ist immer noch ein Teil der Ergebnisse unbefriedigend, da deren relative Wurzel aus dem dem mittleren quadratischen Fehler jenseits der 15 % Marke liegt.

In der amtlichen Statistik herrscht momentan noch eine gewisse Skepsis, diese Methoden auch wirklich zu verwenden, sodass die klassische Sichtweise, eine (asymptotisch) unverzerrte Schätzfunktion unter Berücksichtigung des Stichprobendesigns zu verwenden, oft vorgezogen wird. Das liegt unter anderem daran, dass Formulierungen wie *given the model is true...* abschreckend wirken. Auch sind die Annahmen die bei solchen Modellen getätigt werden nicht immer erfüllt. Dennoch können Methoden der Small Area-Schätzung auch heute schon in „vielen Situationen zu erheblich besseren Schätzungen beitragen“ [Münnich et al., 2013, S. 187]. Trotz der starken Diskussion um die Anwendung solcher Methoden sind [Münnich et al., 2013] und auch ich selbst davon überzeugt, dass diese Methoden in Zukunft immer weiter an Bedeutung gewinnen werden und dementsprechend auch mehr zum Einsatz kommen werden.

So sind Verfahren der Small Area-Schätzung nicht nur nicht nur auf stetige bzw. metrische Zielvariablen anwendbar, sondern lassen sich auch auf kategoriale Zielvariablen erweitern. Das Thema Small Area-Schätzung ist noch lange nicht genug erforscht worden, sodass uns in den nächsten Jahren voraussichtlich noch weitere neue Erkenntnisse und Erweiterungen zu diesem Themengebiet geben wird.

Literatur

- [Battese et al., 1988] Battese, G., Harter, R., and WA, F. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *J Am Stat Assoc*, 83:28–36.
- [Bayrische Staatsregierung, 2016] Bayrische Staatsregierung, D. (2016). *Bezirke, Landkreise, Städte und Gemeinden*. Freistaat Bayern. [<http://www.bayern.de/freistaat/staat-und-kommunen/bezirke-landkreise-staedte-und-gemeinden/> - Letzter Aufruf: 4. Juni 2016].
- [Dieterle, 2010] Dieterle, M. (2010). *Schätzung regionaler Daten mithilfe von Small Area-Schätzmethoden*. Statistisches Bundesamt, Wiesbaden. [<https://www.destatis.de/DE/Publikationen/WirtschaftStatistik/LandForstwirtschaft/SchaetzungRegionalerDaten.pdf> - Letzter Aufruf: 4. Juni 2016].
- [Drew et al., 1982] Drew, J., Singh, M., and Chaudhry, G. (1982). Evaluation of small area estimation techniques for the canadian labour force survey. *Surv Methodol*, 8:17–47.
- [Fay and Herriot, 1979] Fay, R. and Herriot, R. (1979). Estimates of income for small places: an application of james-stein procedures to census data. *J Am Stat Assoc*, 74:269–277.
- [Kauermann and Küchenhoff, 2010] Kauermann, G. and Küchenhoff, H. (2010). *Stichproben, Methoden und praktische Umsetzung in R*. Springer.
- [Münnich et al., 2013] Münnich, R., Burgard, J. P., and Vogt, M. (2013). Small area-statistik; methoden und anwendungen. *ASTA Wirtschafts- und Sozialstatistisches Archiv*, 6:149–191.