

Bericht

Kanton Appenzell-Ausserhoden
Finanzdepartement
Amt für Finanzen

Wirtschaft
Institut für Betriebs- und
Regionalökonomie IBR

Roland Fischer

Dozent

T direkt +41 41 228 99 43
roland.fischer@hslu.ch

Luzern, 4. Mai 2023

Zeitgemässer Finanzausgleich Appenzell Ausserrhoden
Zusatzbericht - Hauptkomponentenanalyse

Roland Fischer
Luzern, 09.05.2023

1.1. Ziel und Funktionsweise der Hauptkomponentenanalyse

1.1.1. Grundlagen und Voraussetzungen

Die Hauptkomponentenanalyse ist ein anerkanntes statistisches Verfahren zur Strukturierung von grossen Datensätzen. Beim Finanzausgleich zwischen und Bund und Kantonen wird z.B. der soziodemografische Lastenausgleich mit Hilfe einer Hauptkomponentenanalyse berechnet. Mit der Hauptkomponentenanalyse werden die Informationen aus einer Vielzahl von Indikatoren auf wenige Indizes oder sogar nur einen Index reduziert. Diese werden als Hauptkomponenten bezeichnet. Sie sollen aber einen möglichst grossen Anteil der in den ursprünglichen Indikatoren vorhandenen Informationen abbilden. Für den GLA und den SLA im Kanton Appenzell-Ausserrhoden werden die Informationen aus sieben verschiedenen strukturellen Indikatoren auf einen Lastenindex reduziert. Der Index soll aber trotz der Reduktion strukturell bedingte Kostenunterschiede zwischen den Gemeinden möglichst gut und plausibel abbilden und somit als zweckmässige Grundlage für den Lastenausgleich dienen.

Damit eine Hauptkomponentenanalyse für den Lastenausgleich herangezogen werden kann, sind folgende zwei Voraussetzungen für die strukturellen Indikatoren wichtig:

- 1) Die Indikatoren sind zumindest teilweise miteinander korreliert, d.h. die Werte der verschiedenen Indikatoren in den Gemeinden weisen einen gewissen Zusammenhang auf. Besteht keine Korrelation zwischen den Indikatoren, so ist es nicht möglich, mit einer Hauptkomponentenanalyse die Informationen aus den einzelnen Indikatoren zusammenzufassen und z.B. auf einen Lastenindex zu reduzieren.
- 2) Die Indikatoren weisen eine gewisse Kostenrelevanz auf, d.h. es besteht z.B. ein Zusammenhang zwischen den Indikatorwerten und dem Nettoaufwand in verschiedenen Aufgabenbereichen der Gemeinden. Nur in diesem Fall ist es plausibel, dass der aus den Indikatoren gebildete Lastenindex dazu geeignet ist, strukturell bedingte Sonderlasten abzubilden.

Beide Voraussetzungen sind im Kanton Appenzell-Ausserrhoden gegeben. Tabelle 1 zeigt die Korrelationskoeffizienten zwischen den strukturellen Indikatoren. Es ist ersichtlich, dass jeder Indikator zumindest mit einem anderen Indikator eine positive oder negative Korrelation aufweist. So korreliert z.B. eine hohe Einwohnerzahl mit einer hohen Bevölkerungsdichte, Altersquote und Sozialhilfequote. Die Verkehrsfläche pro Einwohner:in ist negativ mit der Einwohnerzahl und der Bevölkerungsdichte korreliert, hingegen besteht eine positive Korrelation mit der Jugendquote.

Tabelle 1 Korrelation zwischen den Strukturellen Indikatoren 2023

	Einwohnerzahl, logarithmiert	Durchschnitt. Höhe der Siedlungsfläche (m.ü.M.)	Bevölkerungsdichte	Verkehrsfläche pro Einwohner:in	Jugendquote (Prozent)	Altersquote (Prozent)	Sozialhilfequote
Einwohnerzahl, logarithmiert	1.00						
Durchschnitt. Höhe der Siedlungsfläche (m.ü.M.)	-0.03	1.00					
Bevölkerungsdichte	0.68	-0.39	1.00				
Verkehrsfläche pro Einwohner:in	-0.41	0.13	-0.70	1.00			
Jugendquote (Prozent)	-0.20	0.35	-0.40	0.54	1.00		
Altersquote (Prozent)	0.40	0.19	0.20	-0.27	-0.58	1.00	
Sozialhilfequote	0.66	-0.09	0.53	-0.42	-0.14	0.18	1.00

Durchschnittswerte der Jahre 2019-2021

Datenquellen: Bundesamt für Statistik, Swisstopo

Bereits die Analyse der Korrelationskoeffizienten deutet auf eine Zweiteilung von strukturellen Eigenschaften hin. Auf der einen Seite stehen die grösseren (Zentrums-)Gemeinden, welche eine hohe Bevölkerungsdichte, eine hohe Altersquote und eine hohe Sozialhilfequote aufweisen. Auf der anderen Seite stehen die kleineren, periphereren Gemeinden. Sie zeigen eine tiefe Einwohnerzahl, eine kleine Bevölkerungsdichte, eine hohe Jugendquote, eine Höhenlage und eine grosse Verkehrsfläche pro Einwohner:in.

Tabelle 2 Korrelation zwischen den Strukturellen Indikatoren und dem Nettoaufwand

	Einwohnerzahl, logarithmiert	Durchschnitt. Höhe der Siedlungsfläche (m.ü.M.)	Bevölkerungs- dichte	Verkehrsfläche pro Einwohner:in	Jugendquote (Prozent)	Altersquote (Prozent)	Sozialhilfequote
Allgemeine Verwaltung	-0.48	-0.11	-0.37	0.37	0.01	-0.09	-0.37
Öff. Ordnung und Sicherheit	0.21	0.03	0.07	0.03	-0.09	0.43	0.28
Bildung	-0.47	0.09	-0.33	0.45	0.69	-0.52	-0.19
Kultur, Sport und Freizeit; Kirche	0.72	0.03	0.52	-0.40	-0.23	0.49	0.33
Gesundheit	0.13	0.21	0.08	-0.11	-0.40	0.27	0.10
Soziale Sicherheit	0.67	-0.11	0.46	-0.42	-0.13	0.21	0.74
Umweltschutz und Raumordnung	0.24	-0.15	0.18	0.18	0.02	0.13	0.08
Verkehr	0.82	0.11	0.56	-0.48	-0.25	0.54	0.55
Volkswirtschaft	0.17	-0.04	0.08	-0.09	-0.09	0.27	0.03

Durchschnittswerte der Jahre 2019-2021

Datenquellen: Bundesamt für Statistik, Swisstopo, Gemeindefinanzstatistik AR

Tabelle 2 zeigt die Korrelation zwischen den strukturellen Indikatoren und dem Nettoaufwand in den verschiedenen Aufgabenbereichen der Gemeinden. Es zeigt sich, dass fast alle Indikatoren eine hohe Korrelation mit dem Nettoaufwand in mindestens einem Aufgabenbereich aufweisen. Eine Ausnahme bildet die durchschnittliche Höhe der Siedlungsfläche, die nur schwach mit dem Nettoaufwand in den einzelnen Aufgabenbereichen korreliert ist. Somit weisen fast alle Indikatoren eine Kostenrelevanz auf, womit auch die zweite oben formulierte Voraussetzung für die Anwendung einer Hauptkomponentenanalyse erfüllt ist.

Die Indikatoren fließen in standardisierter Form in die Hauptkomponentenanalyse ein. Das bedeutet, dass sie so transformiert werden, dass sie einen Mittelwert von 0 und eine Standardabweichung von 1 aufweisen. Die Standardisierung ist notwendig, weil die Indikatoren unterschiedliche Messgrößen und Dimensionen aufweisen. Während z.B. die Jugendquote, die Altersquote und die Sozialhilfequote Verhältnisse oder Prozentzahlen darstellen, wird die Verkehrsfläche in Hektaren gemessen. Durch die Standardisierung werden die Indikatoren vergleichbar. Tabelle 3 zeigt die Indikatorwerte für das Ausgleichsjahr 2023, Tabelle 4 die entsprechenden standardisierten Werte. Ein positiver standardisierter Wert bedeutet ein im Vergleich zum Mittelwert der Gemeinden überdurchschnittlich hoher Wert des Indikators, ein negativer standardisierter Wert dementsprechend ein unterdurchschnittlicher Wert.

Tabelle 3 Indikatorwerte 2023

Indikatorwerte der Gemeinden

	Einwohner:in	Log. Einwohner:in	Durchschnitt. Höhe der Siedlungsfläche	Bevölkerungs- dichte (ha)	Jugendquote (in Prozent)	Altersquote (in Prozent)	Sozialhilfequote (in Prozent)	Verkehrsfläche pro Einwohner:in (ha)
Bühler	1'855	7.5	840	3.3	18.6	4.5	1.9	0.007
Gais	3'077	8.0	943	1.5	16.8	5.9	1.3	0.012
Grub	994	6.9	822	2.3	15.7	5.7	1.8	0.009
Heiden	4'206	8.3	800	5.6	15.8	6.6	2.0	0.007
Herisau	15'697	9.7	773	6.2	15.6	5.3	4.5	0.007
Hundwil	960	6.9	794	0.4	23.3	3.4	0.8	0.035
Lutzenberg	1'290	7.2	651	5.7	16.1	2.8	1.1	0.010
Rehetobel	1'730	7.5	936	2.6	15.3	5.5	1.3	0.010
Reute	696	6.5	802	1.4	11.4	7.3	0.9	0.019
Schönengrund	530	6.3	850	1.0	20.0	4.0	1.1	0.015
Schwellbrunn	1'546	7.3	965	0.9	21.2	4.4	0.8	0.021
Speicher	4'422	8.4	906	5.4	18.1	5.9	1.6	0.009
Stein	1'391	7.2	823	1.5	18.6	4.1	0.5	0.014
Teufen	6'385	8.8	838	4.2	15.3	6.8	1.2	0.012
Trogen	1'814	7.5	899	1.8	19.3	4.8	2.8	0.012
Urnäsch	2'275	7.7	845	0.5	18.6	5.4	1.5	0.028
Wald	882	6.8	942	1.3	17.3	3.8	1.0	0.017
Waldstatt	1'810	7.5	820	2.7	17.7	5.5	1.1	0.010
Walzenhausen	2'031	7.6	707	2.9	13.3	5.0	1.3	0.009
Wolfhalden	1'858	7.5	689	2.7	15.7	4.5	1.4	0.012
Mittelwert	2'772	7.6	832	2.7	17.2	5.1	1.5	0.014
Standardabweichung	3'281	0.8	83	1.8	2.6	1.1	0.8	0.007

Durchschnittswerte der Jahre 2019-2021

Datenquellen: BFS, Swisstopo

Tabelle 4 Standardisierte Indikatorwerte 2023

standardisierte Indikatorwerte der Gemeinden

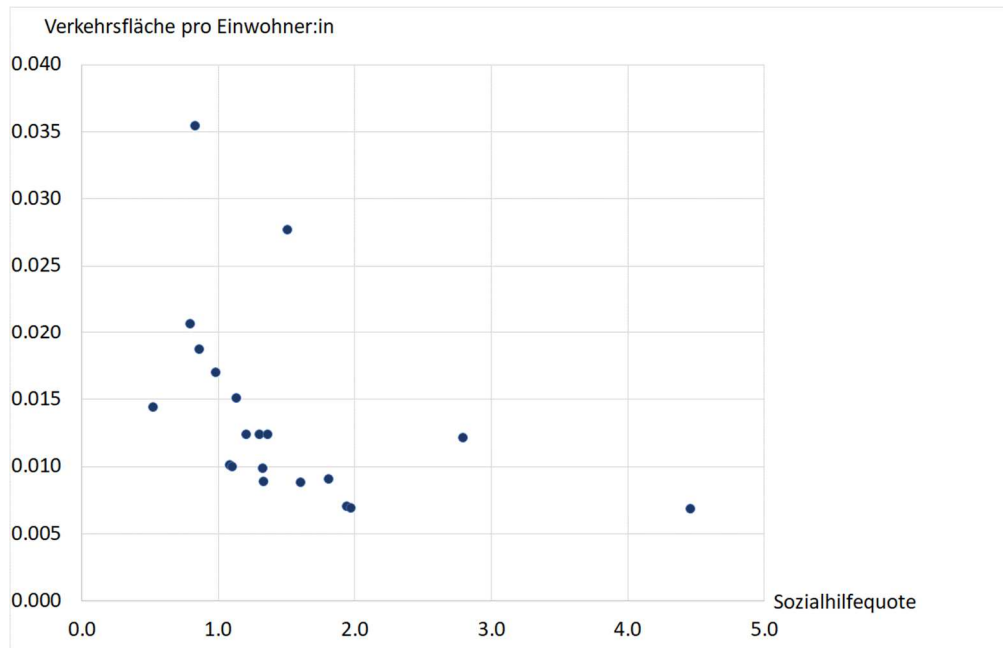
	Log. Einwohner:in	Durch. Höhe der Siedlungsfläche	Bevölkerungsdichte (ha)	Jugendquote	Altersquote	Sozialhilfequote	Verkehrsfläche pro Einwohner:in (ha)
Bühler	0.0	0.1	0.4	0.5	-0.5	0.5	-1.0
Gais	0.6	1.3	-0.7	-0.1	0.7	-0.2	-0.2
Grub	-0.8	-0.1	-0.2	-0.6	0.6	0.4	-0.7
Heiden	1.0	-0.4	1.6	-0.5	1.4	0.6	-1.0
Herisau	2.7	-0.7	2.0	-0.6	0.2	3.5	-1.0
Hundwil	-0.9	-0.5	-1.3	2.3	-1.5	-0.8	3.0
Lutzenberg	-0.5	-2.2	1.7	-0.4	-2.0	-0.5	-0.5
Rehetobel	-0.1	1.2	-0.1	-0.7	0.4	-0.2	-0.6
Reute	-1.3	-0.4	-0.7	-2.2	2.0	-0.8	0.7
Schönengrund	-1.7	0.2	-0.9	1.1	-0.9	-0.4	0.2
Schwellbrunn	-0.3	1.6	-1.0	1.5	-0.6	-0.8	1.0
Speicher	1.1	0.9	1.5	0.4	0.7	0.1	-0.7
Stein	-0.4	-0.1	-0.7	0.5	-0.8	-1.2	0.1
Teufen	1.6	0.1	0.8	-0.7	1.5	-0.3	-0.2
Trogen	-0.1	0.8	-0.5	0.8	-0.2	1.5	-0.2
Urnäsch	0.2	0.1	-1.2	0.5	0.3	0.0	2.0
Wald	-1.0	1.3	-0.8	0.0	-1.1	-0.6	0.5
Waldstatt	-0.1	-0.2	0.0	0.2	0.4	-0.5	-0.5
Walzenhausen	0.1	-1.5	0.1	-1.5	0.0	-0.2	-0.7
Wolfhalden	0.0	-1.7	0.0	-0.6	-0.5	-0.2	-0.2
Mittelwert	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Standardabweichung	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Datenquellen: BFS, Swisstopo

1.1.2. Einfaches Beispiel: geometrische Herleitung

Die Funktionsweise der Hauptkomponentenanalyse kann an einem einfachen, hypothetischen Beispiel illustriert werden: Nehmen wir an, wir würden einen Lastenindex aus nur zwei strukturellen Indikatoren bilden, z.B. mit der Verkehrsfläche pro Einwohnerin und der Sozialhilfequote. Die beiden Indikatoren sind negativ korreliert mit einem Korrelationskoeffizienten von -0.42 (vgl. Tabelle 1). Zudem weisen beide Indikatoren eine Kostenrelevanz auf (vgl. Tabelle 2). Die Verkehrsfläche ist positiv korreliert mit dem Nettoaufwand in der Allgemeinen Verwaltung, der Bildung und dem Bereich Umweltschutz und Raumordnung, die Sozialhilfequote insbesondere mit dem Nettoaufwand für die Soziale Sicherheit. Somit sind die beiden Indikatoren geeignet für eine Hauptkomponentenanalyse. Das Ziel besteht darin, die beiden Indikatoren zu einem Lastenindex zusammenzufassen.

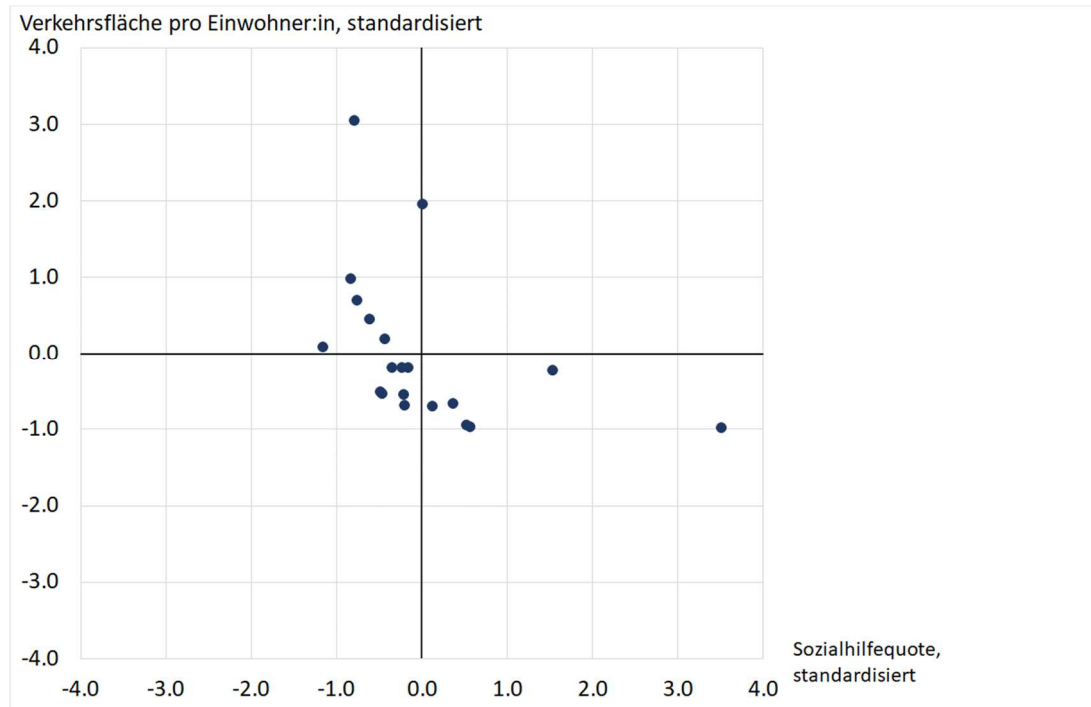
Abbildung 1 Indikatoren Verkehrsfläche pro Einwohner:in und Sozialhilfequote 2023



Datenquelle: BFS

Die Indikatorwerte der Gemeinden sind in Abbildung 1 dargestellt. Klar ersichtlich ist die negative Korrelation zwischen den beiden Indikatoren. Gemeinden mit einer relativ tiefen Sozialhilfequote haben in der Regel eine höhere Verkehrsfläche pro Einwohner:in als Gemeinden mit einer hohen Sozialhilfequote. In einem nächsten Schritt werden die Daten der beiden Indikatoren standardisiert, so dass sie einen Mittelwert von 0 und eine Standardabweichung von 1 aufweisen. Damit werden sie vergleichbar. Die standardisierten Werte der Indikatoren sind in Abbildung 2 dargestellt. Die Korrelation zwischen den beiden Indikatoren wird durch die Standardisierung nicht verändert.

Abbildung 2 Indikatoren Verkehrsfläche pro Einwohner:in und Sozialhilfequote 2023, standardisiert

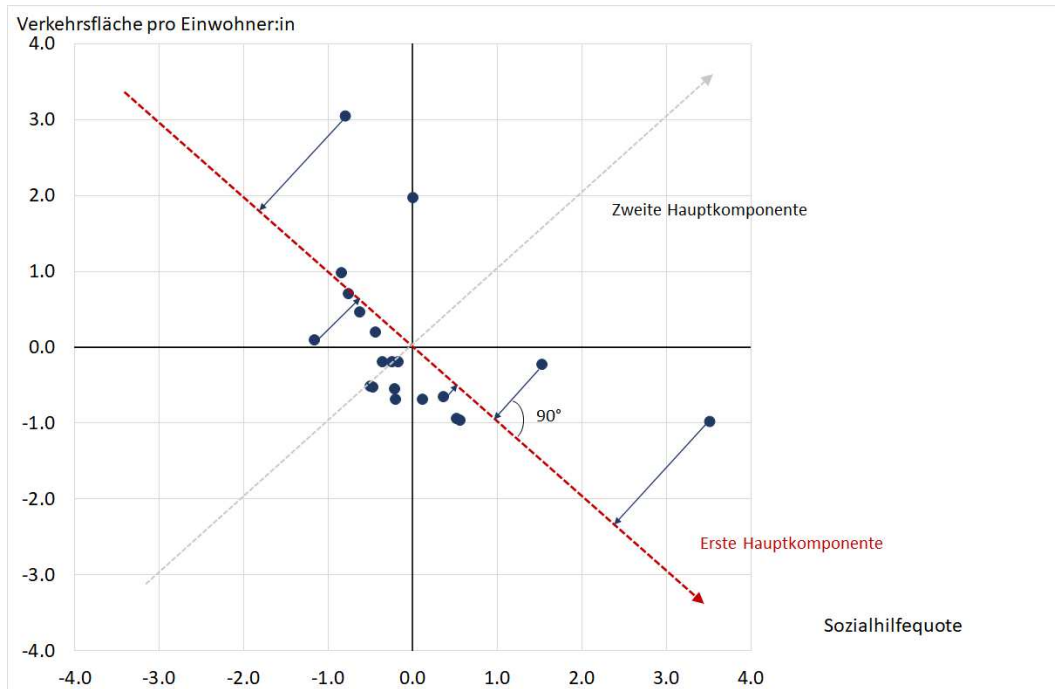


Datenquelle: BFS

Mit der Hauptkomponentenanalyse wird in einem nächsten Schritt diejenige lineare Kombination der beiden Indikatoren gesucht, welche die Informationen der beiden Indikatoren so vollständig wie möglich umfasst. Statistisch ausgedrückt handelt es sich um diejenige lineare Kombination mit dem höchstmöglichen Anteil an der gesamten Varianz der Datenmenge. Bei zwei Indikatoren kann dies sehr gut graphisch dargestellt werden. Es geht darum, in Abbildung 2 eine Gerade in die Punktwolke hineinzulegen, auf welche die Datenpunkte senkrecht, d.h. mit einem Winkel von 90 Grad, projiziert werden. Die Gerade wird so ausgerichtet, dass sie die Ausdehnung der Punktwolke möglichst gut widerspiegelt. Konkret bedeutet dies, dass die Summe der quadratischen Abstände zwischen den Punkten und ihrer Projektion auf der Gerade möglichst klein ausfällt. In unserem Beispiel und aufgrund der Standardisierung der Daten ergibt dies eine Gerade (bzw. ein Vektor), die in einem 45 Grad Winkel vom Quadranten links oben durch den Nullpunkt in den Quadranten rechts unten verläuft. Diese Gerade repräsentiert die erste Hauptkomponente und den aus den beiden Indikatoren gebildete Lastenindex (vgl. Abbildung 3). Die Lastenindexwerte der Gemeinden sind gleich der Distanz zwischen den senkrecht auf die Gerade projizierten Punkte und dem Nullpunkt der ersten Hauptkomponente.

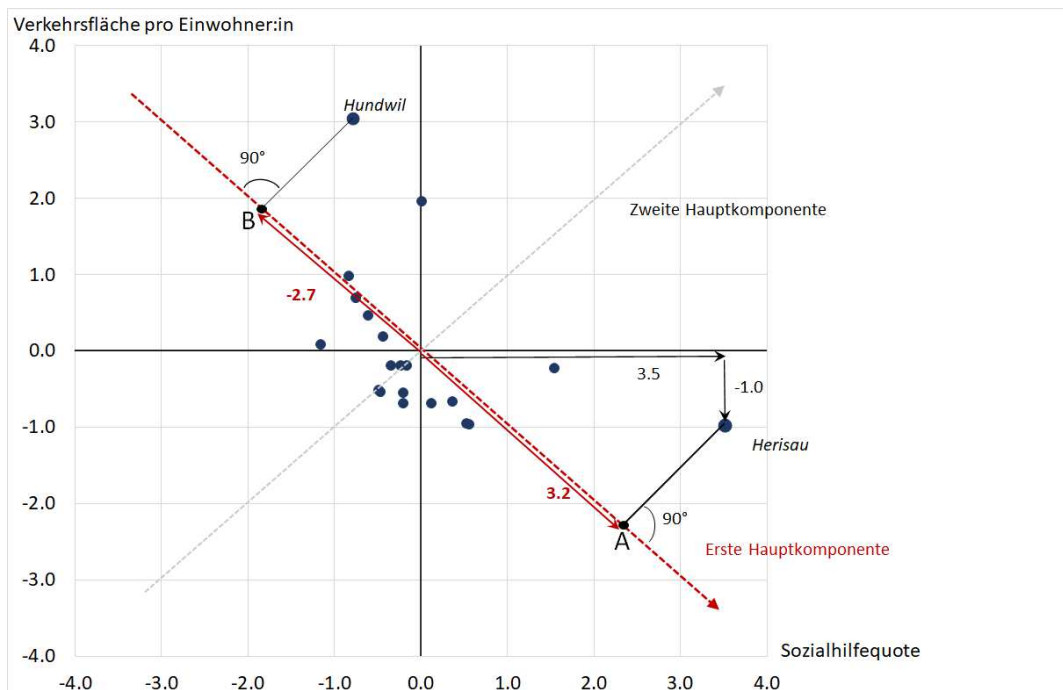
Im rechten Winkel zur ersten Hauptkomponente wird die zweite Hauptkomponente gelegt. Sie repräsentiert den restlichen, nicht in der ersten Komponente enthaltenen, kleineren Anteil an der Varianz der Datenmenge bzw. den darin enthaltenen Informationen. Die Anzahl der Komponenten entspricht stets der Anzahl Indikatoren in einem Datensatz. Diejenigen linearen Kombinationen der Indikatorwerte, welche den grössten Anteil der Varianz der Daten widerspiegeln, werden in absteigender Reihenfolge als erste, zweite, dritte usw. Hauptkomponenten bezeichnet.

Abbildung 3 Hauptkomponenten



Datenquelle: BFS

Abbildung 4 Beispiel: Lastenindex-Werte der Gemeinden Herisau und Hundwil.



Datenquelle: BFS

Abbildung 4 zeigt die graphische Ermittlung der Lastenindex-Werte der Gemeinden Herisau und Hundwil. Herisau weist eine standardisierte Sozialhilfequote von 3.5 und eine standardisierte Verkehrsfläche pro Einwohner:in von -1.0 auf. Diese Werte definieren die Lage von Herisau im Koordinatensystem der beiden Indikatoren. Die senkrechte Projektion Herisaus auf die erste Hauptkomponente ergibt den Punkt A. Der Lastenindex-Wert Herisaus entspricht der Distanz auf dieser Geraden zwischen dem Nullpunkt und dem Punkt A, sie beträgt 3.2. Analog lässt sich der Lastenindex-Wert von Hundwil bestimmen. Die senkrechte Projektion des Datenpunktes von Hundwil ergibt den Punkt B mit einer Distanz vom Nullpunkt von -2.7.

1.1.3. Einfaches Beispiel: algebraische Herleitung

In der Praxis wird die Hauptkomponentenanalyse algebraisch mit Hilfe von Statistik-Software durchgeführt. Wie bereits erwähnt sind die Hauptkomponenten bzw. die Lastenindizes lineare Kombinationen der Indikatoren. Der Lastenindex L_g einer Gemeinde g sei gegeben durch folgende Gleichung:

$$L_g = \mu_V V_g + \mu_S S_g$$

Dabei ist μ_V das Gewicht für den Wert des Indikators Verkehrsfläche einer Gemeinde g , dargestellt durch die Variable V_g , und entsprechend μ_S das Gewicht für den Wert des Indikators Sozialhilfequote einer Gemeinde g , dargestellt durch die Variable S_g . Die beiden Gewichte sind identisch mit dem Eigenvektor des höchsten Eigenwerts der Korrelationsmatrix der beiden standardisierten Indikatoren. Sie ist in Tabelle 5 ersichtlich.

Tabelle 5 Korrelationsmatrix der Indikatoren

	Verkehrsfläche	Sozialhilfequote
Verkehrsfläche	1.00	-0.42
Sozialhilfequote	-0.42	1.00

In algebraischer Form kann diese Matrix wie folgt geschrieben werden:

$$A = \begin{bmatrix} 1 & -0.42 \\ -0.42 & 1 \end{bmatrix}$$

Die Eigenwerte und die Eigenvektoren können für eine 2x2-Matrix durch manuelles Rechnen ermittelt werden. Für grössere Matrizen müssen sie jedoch durch iterative Verfahren geschätzt werden. Die Eigenwerte der Korrelationsmatrix sind

$$\lambda_1 = 1.42$$

$$\lambda_2 = 0.58$$

Die zu den beiden Eigenwerten gehörenden Eigenvektoren sind

$$X_1 = \begin{bmatrix} X_{1,V} \\ X_{1,S} \end{bmatrix} = \begin{bmatrix} -0.71 \\ 0.71 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} X_{2,V} \\ X_{2,S} \end{bmatrix} = \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix}$$

Die Gewichte für den Lastenindex sind gleich den Werten desjenigen Eigenvektors, der aus dem grössten Eigenwert bestimmt wurde. Somit gilt:

$$\begin{bmatrix} \mu_V \\ \mu_S \end{bmatrix} = \begin{bmatrix} X_{1,V} \\ X_{1,S} \end{bmatrix}$$

$$\begin{bmatrix} \mu_V \\ \mu_S \end{bmatrix} = \begin{bmatrix} -0.71 \\ 0.71 \end{bmatrix}$$

Daraus resultieren in unserem Beispiel folgende Gewichte.

$$\mu_V = -0.71$$

$$\mu_S = 0.71$$

Der Lastenindex einer Gemeinde g wird durch Multiplikation dieser Gewichte mit den Werten der Indikatoren Verkehrsfläche und Sozialhilfequote berechnet:

$$L_g = -0.71 \cdot V_g + 0.71 \cdot S_g$$

Der so ermittelte Lastenindex entspricht in der Terminologie der Hauptkomponentenanalyse dem Eigenvektor der ersten Hauptkomponente. Die Lastenindexwerte der Gemeinden sind in Tabelle 6 dargestellt.

Für die Gemeinde Herisau zum Beispiel lautet die Berechnung ihres Lastenindexwertes wie folgt:

$$L_{Herisau} = -0.71 \cdot -1.0 + 0.71 \cdot 3.5 = 3.2$$

Unter Berücksichtigung der Anforderung, dass nur Sonderlasten ausgeglichen werden, wird mit den positiven Werten dieses Index ein Indikator Sonderlasten definiert. Ausgleichszahlungen erhalten alle

Gemeinden, die positive Indikatorwerte aufweisen. Die Gemeinde Herisau würde somit Ausgleichszahlungen erhalten.

Für die Gemeinde Hundwil als weiteres Beispiel lautet die Berechnung wie folgt:

$$L_{\text{Hundwil}} = -0.71 \cdot 3.0 + 0.71 \cdot -0.8 = -2.7$$

Die Gemeinde Hundwil würde bei diesem Lastenindex keine Ausgleichszahlungen erhalten, da ihr Lastenindex einen negativen Wert aufweist. Aufgrund der Ausprägungen der Indexwerte könnte dieser Lastenindex zum Beispiel als Zentrallastenindex interpretiert werden.

Tabelle 6 Einfaches Beispiel eines Lastenindex

	Verkehrsfläche pro Einwohner:in	Sozialhilf- equote	Verkehrsfläche standardisiert	Sozialhilfequote standardisiert	Lastenindex	Indikator Sonderlasten
	ha	Prozent	Gew. = -0.71	Gew. = 0.71		
Bühler	0.007	1.9	-1.0	0.5	1.0	1.0
Gais	0.012	1.3	-0.2	-0.2	0.0	
Grub	0.009	1.8	-0.7	0.4	0.7	0.7
Heiden	0.007	2.0	-1.0	0.6	1.1	1.1
Herisau	0.007	4.5	-1.0	3.5	3.2	3.2
Hundwil	0.035	0.8	3.0	-0.8	-2.7	
Lutzenberg	0.010	1.1	-0.5	-0.5	0.0	0.0
Rehetobel	0.010	1.3	-0.6	-0.2	0.2	0.2
Reute	0.019	0.9	0.7	-0.8	-1.0	
Schönengrund	0.015	1.1	0.2	-0.4	-0.4	
Schwellbrunn	0.021	0.8	1.0	-0.8	-1.3	
Speicher	0.009	1.6	-0.7	0.1	0.6	0.6
Stein	0.014	0.5	0.1	-1.2	-0.9	
Teufen	0.012	1.2	-0.2	-0.3	-0.1	
Trogen	0.012	2.8	-0.2	1.5	1.3	1.3
Urnäsch	0.028	1.5	2.0	0.0	-1.4	
Wald	0.017	1.0	0.5	-0.6	-0.8	
Waldstatt	0.010	1.1	-0.5	-0.5	0.1	0.1
Walzenhausen	0.009	1.3	-0.7	-0.2	0.3	0.3
Wolfhalden	0.012	1.4	-0.2	-0.2	0.0	0.0
Mittelwert	0.014	1.5	-0.0	0.0		
Standardabweichung	0.007	0.8	1.0	1.1		

Datenquelle: BFS

Zu beachten ist, dass es sich hier um eine vereinfachte, nicht repräsentative Hauptkomponentenanalyse handelt. In der Praxis ist eine Hauptkomponentenanalyse mit nur zwei Variablen von geringem Nutzen, da der Nutzen der Datenreduktion von zwei auf einen Indikator sehr klein ist. Zudem können in einem System, das nur aus zwei Indikatoren besteht, die Sonderlasten nicht umfassend abgebildet werden. Die Beschränkung auf zwei Variablen erlaubt jedoch eine Veranschaulichung des Konzepts der Hauptkomponentenanalyse.

Weiter ist zu beachten, dass eine Hauptkomponentenanalyse im Hinblick auf einen Lastenausgleich nicht immer plausible und verwertbare Ergebnisse hervorbringt, selbst wenn die Anforderungen für die Korrelation der Indikatoren und der Kostenrelevanz erfüllt sind. Die Tauglichkeit der Methode kann erst beim Vorliegen der Eigenvektoren bzw. der Hauptkomponenten abschliessend beurteilt werden.

1.2. Aktualisierung der Berechnung des soziodemografischen und geografisch-topografischen Lastenindex

1.2.1. Ergebnis der Hauptkomponentenanalyse und Interpretation

Die Korrelationsmatrix für die Hauptkomponentenanalyse ist bereits weiter vorne in Tabelle 1 dargestellt. Deren Eigenwerte sind in Tabelle 7 ersichtlich. Die Eigenwerte sind in absteigender Reihenfolge aufgelistet. Sie zeigen, welchen Anteil der gesamten Varianz der Indikatoren bzw. der in den Indikatoren enthaltenden Informationen die Komponenten beinhalten. Die Daten zeigen, dass der erste Eigenwert einen Anteil von 46 Prozent an der Summe der Eigenwerte hat. Das bedeutet, dass ein Lastenindex bestehend aus der ersten Hauptkomponente 46 Prozent der Informationen der sieben Indikatoren beinhaltet. Obwohl also die sieben Indikatoren auf einen einzigen Index reduziert werden, bildet dieser Index fast die Hälfte der mit den sieben Indikatoren gemessenen strukturellen Eigenschaften der Gemeinden ab.

Tabelle 7 Hauptkomponentenanalyse: Eigenwerte der Korrelationsmatrix

Nummer	Eigenwert	Anteil	Kumulierter Wert	Kumulierter Anteil
1	3.21	0.46	3.21	0.46
2	1.27	0.18	4.48	0.64
3	1.24	0.18	5.72	0.82
4	0.62	0.09	6.33	0.90
5	0.40	0.06	6.74	0.96
6	0.15	0.02	6.89	0.98
7	0.11	0.02	7.00	1.00
Summe	7.00	1.00		

Die zu den Eigenwerten gehörenden Eigenvektoren bzw. Hauptkomponenten sind in Tabelle 8 dargestellt. Die Koeffizienten sind die Gewichte der Indikatoren. Die Analyse der Gewichte der ersten Hauptkomponente zeigt, dass sie sich sehr gut als Indikator für soziodemografische Sonderlasten eignet. So fließen die Einwohnerzahl, die Bevölkerungsdichte, die Altersquote und die Sozialhilfequote mit einem positiven Gewicht in den Indikator ein. Hohe Werte dieser Indikatoren dürften somit in grösseren Gemeinden mit Zentrumslasten zu finden sein. Die weiteren Eigenvektoren zeigen hingegen im Hinblick auf die Abbildung von Sonderlasten keine plausiblen linearen Kombinationen der Indikatoren.

Tabelle 8 Hauptkomponentenanalyse: Eigenvektoren der Korrelationsmatrix

	Eigenvektoren						
	1	2	3	4	5	6	7
Log. Einwohner:in	0.43	0.24	0.35	0.33	0.40	-0.46	-0.40
Durchschnittliche Höhe der Siedlungsfläche	-0.17	0.77	0.12	-0.44	0.02	-0.25	0.33
Bevölkerungsdichte	0.48	-0.22	0.17	-0.13	0.45	0.21	0.65
Verkehrsfläche pro Einwohner:in	-0.44	0.07	0.08	0.74	0.08	-0.14	0.47
Jugendquote	-0.36	0.09	0.63	-0.05	0.25	0.58	-0.24
Altersquote	0.29	0.53	-0.46	0.32	0.07	0.56	-0.08
Sozialhilfequote	0.38	0.11	0.47	0.17	-0.75	0.11	0.14

Zu beachten ist aber, dass die Inverse der ersten Hauptkomponente strukturelle Ursachen von geografisch-topografischen Sonderlasten abbildet (vgl. Tabelle 9). So fließen bei der Inverse die Einwohnerzahl und die Bevölkerungsdichte mit einem negativen Gewicht ein. Eine tiefe Einwohnerzahl ist in vielen Aufgabenbereichen mit höheren Kosten verbunden, da eine kleine Gemeinde Grössenvorteile bei der Bereitstellung von Gütern und Dienstleistungen nicht ausschöpfen kann, zum Beispiel im Bereich der Administration oder der Bildung. Eine tiefe Bevölkerungsdichte entspricht einer dünnen Besiedelung, was mit höheren Kosten für die Verkehrs- und Versorgungsinfrastruktur verbunden ist. Zudem haben beim inversen Indikator die durchschnittliche Höhe der Siedlungsfläche und die Verkehrsfläche pro Einwohner:in ein positives Gewicht. Hohe Werte dieser Indikatoren sind ebenfalls bei kleineren, dünn besiedelten Gemeinden zu erwarten. Die Inverse der ersten Hauptkomponente kann deshalb als guter Indikator für geografisch-topografische Sonderlasten verwendet werden.

Tabelle 9 Hauptkomponentenanalyse: Soziodemografischer und geografisch-topografischer Index

	Gewichte SLA*	Gewichte GLA**
Log. Einwohner:in	0.43	-0.43
Durchschnittliche Höhe der Siedlungsfläche	-0.17	0.17
Bevölkerungsdichte	0.48	-0.48
Verkehrsfläche pro Einwohner:in	-0.44	0.44
Jugendquote	-0.36	0.36
Altersquote	0.29	-0.29
Sozialhilfequote	0.38	-0.38

* = Eigenvektor der ersten Hauptkomponente

** = Inverser Eigenvektor der ersten Hauptkomponente

Die erste Hauptkomponente zeigt somit die unterschiedliche Ausprägung von strukturellen Merkmalen sowohl in den urbaneren Gemeinden als auch in den ländlichen Gemeinden. Die Verwendung der gleichen Hauptkomponente sowohl als SLA- als auch, in der inversen Form, als GLA-Index hat jedoch zur Folge, dass eine Gemeinde nicht gleichzeitig Beitragsleistungen aus beiden Lastenausgleichsgefässen erhalten kann. Andererseits erhält jede Gemeinde entweder geografisch-topografischen oder soziodemografischen Lastenausgleich. Allerdings fallen die Ausgleichszahlungen bei Gemeinden mit Indexwerten nahe bei 0 relativ klein aus.

Tabelle 10 Standardisierte Indikatoren 2023, Gewichte, SLA- und GLA-Index

	Log. Einwohner:in	Durchschn. Höhe der Siedlungs- fläche	Bevölkerungs- dichte	Verkehrs- fläche pro Einwohner:in	Jugend- quotient	Alters- quotient	Sozialhilfe- quote	Index SLA	Index GLA
<i>Gewichte SLA</i>	<i>0.43</i>	<i>-0.17</i>	<i>0.48</i>	<i>-0.44</i>	<i>-0.36</i>	<i>0.29</i>	<i>0.38</i>		
<i>Gewichte GLA</i>	<i>-0.43</i>	<i>0.17</i>	<i>-0.48</i>	<i>0.44</i>	<i>0.36</i>	<i>-0.29</i>	<i>-0.38</i>		
Bühler	-0.0	0.1	0.4	-1.0	0.5	-0.5	0.5	0.4	-0.4
Gais	0.6	1.3	-0.7	-0.2	-0.1	0.7	-0.2	0.0	0.0
Grub	-0.8	-0.1	-0.2	-0.7	-0.6	0.6	0.4	0.4	-0.4
Heiden	1.0	-0.4	1.6	-1.0	-0.5	1.4	0.6	2.5	-2.5
Herisau	2.7	-0.7	2.0	-1.0	-0.6	0.2	3.5	4.3	-4.3
Hundwil	-0.9	-0.5	-1.3	3.0	2.3	-1.5	-0.8	-3.8	3.8
Lutzenberg	-0.5	-2.2	1.7	-0.5	-0.4	-2.0	-0.5	0.6	-0.6
Rehetobel	-0.1	1.2	-0.1	-0.6	-0.7	0.4	-0.2	0.2	-0.2
Reute	-1.3	-0.4	-0.7	0.7	-2.2	2.0	-0.8	-0.1	0.1
Schönengrund	-1.7	0.2	-0.9	0.2	1.1	-0.9	-0.4	-2.1	2.1
Schwellbrunn	-0.3	1.6	-1.0	1.0	1.5	-0.6	-0.8	-2.4	2.4
Speicher	1.1	0.9	1.5	-0.7	0.4	0.7	0.1	1.5	-1.5
Stein	-0.4	-0.1	-0.7	0.1	0.5	-0.8	-1.2	-1.4	1.4
Teufen	1.6	0.1	0.8	-0.2	-0.7	1.5	-0.3	1.7	-1.7
Trogen	-0.1	0.8	-0.5	-0.2	0.8	-0.2	1.5	-0.1	0.1
Urnäsch	0.2	0.1	-1.2	2.0	0.5	0.3	0.0	-1.5	1.5
Wald	-1.0	1.3	-0.8	0.5	0.0	-1.1	-0.6	-1.8	1.8
Waldstatt	-0.1	-0.2	0.0	-0.5	0.2	0.4	-0.5	0.1	-0.1
Walzenhausen	0.1	-1.5	0.1	-0.7	-1.5	-0.0	-0.2	1.1	-1.1
Wolfhalden	-0.0	-1.7	-0.0	-0.2	-0.6	-0.5	-0.2	0.4	-0.4

Durchschnittswerte der Jahre 2019-2021

Datenquellen: BFS, Swisstopo

Die standardisierten Indikatorwerte und Lastenindexwerte der Gemeinden sind in Tabelle 10 ersichtlich. In den Kopfzeilen eingefügt sind zudem die Gewichte der einzelnen Indikatoren für die Berechnung des SLA- und des GLA-Index. Zur Berechnung des SLA-Index und des GLA-Index der Gemeinden werden die standardisierten Indikatorwerte mit den jeweiligen Gewichten multipliziert und addiert. Für die Gemeinde Herisau lautet z.B. die Gleichung für den soziodemografischen Lastenindex ($S_{Herisau}$) wie folgt:

$$S_{Herisau} = (0.43 \cdot 2.7) + (-0.17 \cdot -0.7) + (0.48 \cdot 2.0) + (-0.44 \cdot -1.0) + (-0.36 \cdot -0.6) + (0.29 \cdot 0.2) + (0.38 \cdot 3.5) = 4.3$$

Herisau hat im Vergleich zum Mittelwert der Gemeinden eine überdurchschnittlich hohe Einwohnerzahl. Da diese im SLA positiv gewichtet ist, hat sie eine positive Wirkung auf den SLA-Index. Gleiches gilt für die Bevölkerungsdichte und die Sozialhilfequote, wo Herisau ebenfalls überdurchschnittlich hohe Werte aufweist. Bei der Höhenlage, der Verkehrsfläche und der Jugendquote zeigt Herisau hingegen unterdurchschnittliche Werte. Da diese Indikatoren mit einem negativen Gewicht in den SLA-Index einfließen, resultiert für Herisau auch bei ihnen eine positive Wirkung auf ihren SLA-Index. Der SLA-Index der Gemeinde Herisau liegt mit 4.3 Punkten deutlich über dem Durchschnitt der Gemeinden, so dass Herisau hohe Ausgleichszahlungen aus dem SLA erhält. Die Gleichung für den GLA-Index multipliziert die standardisierten Indikatorwerte mit den inversen Gewichten, so dass der GLA-Index von Herisau bei -4.3 Punkten liegt.

Als weiteres Beispiel wird der GLA-Index der Gemeinde Hundwil berechnet:

$$G_{Hum.} = (-0.43 \cdot -0.9) + (0.17 \cdot -0.5) + (-0.48 \cdot -1.3) + (0.44 \cdot 3.0) + (0.36 \cdot 2.3) + (-0.29 \cdot -1.5) + (-0.38 \cdot -0.8) = 3.8$$

Die Einwohnerzahl der Gemeinde Hundwil liegt unter dem Durchschnitt der Gemeinden. Da jedoch dieser Indikator im GLA ein negatives Gewicht besitzt, beeinflusst die Einwohnerzahl den GLA-Index positiv. Das Gleiche gilt für die Bevölkerungsdichte. Ebenfalls eine positive Wirkung haben die hohe Verkehrsfläche pro Einwohner:in sowie der hohe Jugendquotient. Eine leicht negative Wirkung geht hingegen von der Höhenlage aus, da der Indikator leicht unter dem Durchschnitt der Gemeinden liegt. Die tiefe Altersquote und Sozialhilfequote beeinflussen hingegen den GLA-Index positiv, da sie im GLA negative Gewichte aufweisen. Insgesamt beträgt der GLA-Index von Hundwil 3.8 Punkte.

1.2.2. Kostenrelevanz des Lastenindex

Die beiden Lastenindizes sind dazu geeignet, Sonderlasten, d.h. durch strukturelle Faktoren bedingte überdurchschnittlich hohe Ausgaben, abzubilden. Diese Hypothese lässt sich mit Regressionsanalysen überprüfen. Diese haben zum Ziel, einen statistisch signifikanten Zusammenhang zwischen den Werten des Lastenindex und dem gesamten Nettoaufwand der Gemeinden zu schätzen. In der Vernehmlassung wurde das tiefe Bestimmtheitsmass (angepasstes R-Quadrat) der in der Grundkonzeption dargestellten Schätzgleichung kritisiert. Je nach gewählter Spezifikation des Modells werden durch das Schätzmodell nur rund 20 bis 50 Prozent der Aufwandunterschiede zwischen den Gemeinden erklärt.

Tabelle 11 Nettoaufwand der Gemeinden als Funktion von Lastenindex und Nettoertrag im Bereich Finanzen und Steuern*

Variable	Koeffizient	Standard- abweichung	t-Wert	Probabilität
Konstante	1'620	616	2.630	0.018
SLA-Index ² (=GLA-Index ²)	22.65	9.3	2.428	0.027
Nettoertrag Finanzen und Steuern	0.50	0.15	3.290	0.004
R-Quadrat	0.52			
Angepasstes R-Quadrat	0.46			

Ein tiefes R-Quadrat allein lässt jedoch kein abschliessendes Urteil über die Kostenrelevanz des Lastenindex zu. Zum einen zeigen die Schätzgleichungen, dass der positive Quotient für den quadrierten Lastenindex eine statistische Signifikanz auf dem 5-Prozent-Niveau (Probabilität) aufweist. Das bedeutet, dass die Null-Hypothese, wonach der Koeffizient Null beträgt, mit einer Irrtumswahrscheinlichkeit von 5 Prozent verworfen wird. Es ist somit sehr wahrscheinlich, dass Gemeinden mit einem hohen SLA- oder GLA-Index einen überdurchschnittlich hohen Nettoaufwand haben und somit Sonderlasten aufweisen.

Zudem muss beachtet werden, dass überdurchschnittlich hohe Ausgaben in den seltensten Fällen durch strukturelle Faktoren allein bedingt sind. So ist z.B. das Ausgabenniveau von Gemeinden in der Regel auch von der Ertragslage abhängig. Gemeinden mit einem hohen Steuertrag pro Einwohner:in tätigen in der Regel höhere Pro-Kopf-Ausgaben, selbst dann, wenn keine Sonderlasten vorliegen. Zudem ist das Ausgabenniveau massgeblich von den Präferenzen der politischen Entscheidungsträger:innen und der Stimmberechtigten geprägt. Um solchen Faktoren Rechnung zu tragen können so genannte Kontrollvariablen in das Schätzmodell eingefügt werden. Im Bericht zur Grundkonzeption wurde der Ressourcenindex als Kontrollvariable verwendet. In Tabelle 11 wurde in einem leicht angepassten Modell der Nettoertrag im Bereich Finanzen und Steuern herangezogen, was das R-Quadrat des Schätzmodells im Vergleich zur Variante mit dem Ressourcenindex erhöht. Bestätigt wird durch diese Schätzgleichung die statistische Signifikanz des Koeffizienten für den Lastenindex. Keine Kontrollvariable wurde für die politischen Präferenzen verwendet. Sie fliessen somit in den nicht-erklärbaren Teil des Nettoaufwands ein (Residuum).