

Methodenbericht

Zur Methodik der
Tamedia-Abstimmungsumfrage
Stand: August 2017





Zusammenfassung

Die Tamedia-Abstimmungsfrage haben sich als wesentliche Informationsquelle im Vorfeld von eidgenössischen Abstimmungen etabliert. Unsere Resultate basieren auf gewichteten Analysen grosser Onlinestichproben. Die gewichtete Analyse von Onlinestichproben stützt sich, im Vergleich zur Standardmethode der Telefonumfragen, auf andere statistische Methoden. Die Grundlagen der gewichteten Analysen sind noch weniger bekannt als die Theorie der Zufalls-stichprobe, auf welcher Telefonumfragen aufbauen. Dieses Dokument diskutiert die Fallstricke und Vorteile der beiden Verfahren und verweist auf zahlreiche Studien und Blogeinträge für vertiertere Informationen. Der letzte Teil dieses Dokumentes beantwortet oft gestellte Fragen zu unseren Analysen und beschreibt unsere Zusammenarbeit mit Tamedia. Zudem gehen wir auf Fragen der Manipulierbarkeit von Onlineumfragen ein und diskutieren die allgemeine Bedeutung von Umfragen in Volksabstimmungen. Grundsätzlich sollten Umfrageresultate vorsichtig als Indikatoren interpretiert werden. Weitere Fragen können gerne per [Mail](mailto:info@leewas.ch) gestellt werden. Sofern diese von allgemeinem Interesse sind, werden wir sie in überarbeiteten Versionen dieses Dokumentes beantworten.



Ausgangslage

In kurzer Zeit entwickelte sich die gewichtete Tamedia-Abstimmungsfrage zu einer viel beachteten Informationsquelle über die Bevölkerungsmeinung in eidgenössischen Abstimmungskampagnen (siehe Berichte in der *AZ*, *DIE ZEIT* oder *NZZ*). Seit Februar 2014 waren die Tamedia Analysen mindestens so verlässlich wie die etablierten SRG-Trendumfragen (bei mehr als der Hälfte der Vorlagen war die letzte Welle der Tamedia-Abstimmungsfrage näher am Endresultat als die letzte SRG-Trendumfrage, siehe Tabelle 1 auf Seite 8). Trotzdem wird das noch junge Analyseverfahren kritisch hinterfragt. Die Skepsis ist unserer Meinung nach insofern gerechtfertigt, als unsere gewichteten Analysen—wie alle Umfrageverfahren—auf theoretischen Annahmen beruhen und mit Fallstricken in der praktischen Umsetzung verbunden sind. Journalistinnen, Politiker und Bürgerinnen sollten generell Resultate von Meinungsumfragen als Indikatoren, nicht als genaue Momentaufnahmen, interpretieren.

Auch wenn generelle Zurückhaltung mit der Interpretation von Umfrageresultaten richtig ist, seriös ausgewertete Onlineumfragen sind nicht kritischer zu betrachten als Telefonumfragen, wie die Erfahrungen in der Schweiz und im Ausland zeigen. In den USA, wo die Grundlagen der neuen gewichteten Analysemethodik entwickelt wurden, ist eine hitzige Debatte entfacht, nachdem die *New York Times* und *CBS News* bekannt gegeben haben, auf Telefonumfragen zu verzichten und stattdessen mit gewichteten Onlineumfragen von *YouGov* zu arbeiten. Der Präsident der US Gesellschaft für Umfrageforschung, Michael Link, hat die Qualität der *YouGov*-Analysen angezweifelt (siehe [hier](#)), worauf der international renommierte Statistiker Andrew Gelman von der Columbia University die Stellungnahme von Michael Link als innovationsfeindlich und unseriös kritisiert hat (siehe [hier](#) und [hier](#)). Die Methodikdebatte muss nicht in amerikanischer Aufgeregtheit geführt werden. Dieser Beitrag soll im Sinne einer konstruktiven Diskussion interessierten Kreisen helfen die Fallstricke und Vorteile von Telefon- und Onlineumfragen besser zu verstehen. Dazu diskutieren wir die Annahmen, statistischen Grundlagen und Herausforderungen der beiden Verfahren.

Während Telefonumfragen auf der Theorie der Zufallsstichprobe aufbauen, folgen die Gewichtungsverfahren von Onlineumfragen unter anderem den neueren Erkenntnissen der Forschung zu Mehrebenenmodellen und Poststratifikation. Die methodischen Fallstricke der Gewichtungsanalysen sind anders gelagert als die der Standardmethode, aber in der praktischen Umsetzung keineswegs schwerwiegender. Im Gegenteil, zwei wesentliche Fakten sprechen für die neueren Gewichtungsverfahren: über Telefoninterviews repräsentative Stichproben zu generieren wird zunehmend schwieriger und onlinegenerierte Gewichtungsanalysen haben sich als vergleichsweise zuverlässig erwiesen.



Standardmethode: Analyse von Telefonstichproben

Die Standardmethode von Telefonumfragen baut auf der Theorie der Zufallsstichprobe auf. Aus der Stimmbevölkerung wird eine "repräsentative" Stichprobe von 1,000 bis 1,500 Befragten interviewt. Wenn die Stichprobe tatsächlich zufällig generiert wurde—das heisst, alle StimmbürgerInnen haben dieselbe Teilnahmewahrscheinlichkeit—ist die Analyse der Daten vergleichsweise einfach, präzise und robust: Die Unterstützung in der Stichprobe für eine Vorlage kann direkt auf die Stimmbevölkerung extrapoliert werden (wobei noch die Unsicherheit berechnet wird). Die Vorteile der Standardmethode sind die wenigen Annahmen in der Datenauswertung und die elegante theoretische Fundierung.

Die Herausforderung (oder Achillesferse) liegt im Generieren der repräsentativen Zufallsstichprobe. Theoretisch sollte jede Wählerin die gleiche Wahrscheinlichkeit haben, in der Stichprobe interviewt zu werden. In der Praxis ist das wegen Problemen im Abdeckungsgrad und der Nicht-Teilnahme an Umfragen nicht der Fall. Während einige Institute mit Telefonlisten arbeiten (die zum Teil einen relativ kleinen Ausschnitt der Stimmbevölkerung abdecken), gilt das sogenannte "random digitling", wo zufällig Telefonnummern generiert werden, als bevorzugtes Verfahren, um eine möglichst hohe Abdeckung der Stimmbevölkerung zu gewährleisten. Eine hohe Abdeckung kann ebenfalls erreicht werden, wenn die Untersuchungsstichprobe aus der Grundgesamtheit gezogen wird und erst in einem zweiten Schritt die Telefonnummern der ausgewählten Individuen identifiziert werden.

Massgebend ist neben der Auswahl der Stichprobe die Erreichbarkeit und Nicht-Teilnahme (bzw. Verweigerung) der Personen, die kontaktiert werden. In aktuellen Telefonumfragen bleiben bis zu 90% der Anrufe unbeantwortet. Über die Jahre ist die Verweigerung und Nicht-Erreichbarkeit deutlich gestiegen (vor 15 Jahren war die Ausschöpfungsquote noch rund doppelt so gut).¹ Problematisch ist die hohe Nicht-Erreichbarkeit und Verweigerung, weil sich die relativ wenig Personen, welche an den Umfragen teilnehmen, systematisch von der grossen Mehrheit unterscheidet, die auch hätte interviewt werden sollten. Wir haben dazu die VOX-Daten aus den Jahren 2010 – 2012 zusammen mit Patrice Siegrist untersucht (VOX Berichte 102 – 109). In diesen Umfragen wurden systematisch zu viele SP-Sympathisanten (+9%) und Wähler mit Universität Abschluss (+8%) befragt, während die Unterstützer der SVP (–5%) und Personen der tiefsten zwei Bildungskategorien untervertreten waren (–8%; –10%). Diese grossen systematischen Verzerrungen in den Stichproben sind in Fachkreisen bekannt. Hanspeter Kriesi vom European University Institute argumentierte beispielsweise, die zu tiefen Umfragewerte im Vorfeld der Minarettinitiative seien insbesondere durch eine zu geringe Teilnahme von schlechter gebildeten Wählern an der Umfrage zu erklären (siehe diesen [Blogeintrag](#)).

Eine neueres Forschungsergebnis von Andrew Gelman und Kollegen von der Columbia University zeigt zudem, dass die Bereitschaft an Umfragen teilzunehmen damit zu tun hat, ob über die eigene Partei (und ihre Wahlaussichten) medial positiv berichtet wird. Ein Teil der Schwankungen in den Umfrageresultaten hat gemäss dieser Studie nicht mit tatsächlichen Verschiebungen in der Wählerschaft zu tun, sondern mit der politischen

¹Bei der VOX 107 (siehe Technischer Bericht vom 10.7.2012) beispielsweise wurden von 14'331 zufällig generierten und gültigen Telefonnummern 1'507 Personen befragt (Ausschöpfung = 10.5%). In der VOX 72 (siehe Technischer Bericht vom 15.12.2000) war die Ausschöpfungsquote bei 19.5% (1'024 verwendbare Interviews bei 5'244 gültigen Ursprungsnummern) und in der VOX 62 (siehe Technischer Bericht vom 22.10.1997) bei 23% (1'001 verwendbare Interviews bei 4'374 gültigen Ursprungsnummern).



und medial bedingten Teilnahmebereitschaft an Umfragen ([hier ist der Blogeintrag dazu](#); [hier das Papier](#)). Zuletzt stellt sich auch die Frage, ob die Telefoninterviewpartnerinnen, die erreicht werden, die Wahrheit sagen. Dieses Problem wird in der Literatur unter anderem am Beispiel der jeweils ca. 20% zu hohen Stimmbeteiligung in Stichproben von Telefonumfragen diskutiert. Offensichtlich sagen einige Interviewpartner fälschlicherweise, sie hätten an Abstimmungen teilgenommen (soziale Erwünschtheit). Pascal Sciarini und Kollegen von der Université de Genève untersuchen mit hochstehenden Daten, wie stark die systematischen Verzerrungen in der Stimmbeteiligung durch die Nicht-Erreichbarkeit und durch Falschaussagen beeinflusst werden.

Fazit: Die Stärke der Standardmethode liegt in den wenigen Annahmen bezüglich der Auswertung der Daten. Ob die Theorie der Zufallsstichprobe anwendbar ist, hängt von der Stichprobe ab. Die Schwierigkeiten über Telefoninterviews eine repräsentative Zufallsstichprobe zu generieren haben stark zugenommen. Einige Massnahmen, wie zum Beispiel die Berücksichtigung von Natelnummern und die Auswahl der Kontaktpersonen mittels "random digitling", verbessern zwar den Abdeckungsgrad, lösen aber das Hauptproblem nicht – nämlich die steigende Nicht-Erreichbarkeit und Verweigerung.



Gewichtete Analyse von grossen Onlinestichproben

Die gewichtete Analyse von grossen Onlinestichproben folgt nicht der Theorie der Zufallsstichprobe. Anstatt eine relativ kleine Stichprobe auszuwerten—(die über ein Zufallsverfahren hätte zustande kommen sollen)—, werden Verzerrungen in Online-Stichproben akzeptiert. Die Herausforderung liegt darin, die Verzerrungen auszugleichen. Dabei werden politische, geographische und demographische Informationen in einem mehrstufigen Verfahren analysiert. Die theoretische Fundierung der gewichteten Analysen orientiert sich unter anderem an der Forschung zu Mehrebenenmodellen und Poststratifikation (siehe dazu Papiere [hier](#), [hier](#) und [hier](#); und Blogbeitrag [hier](#)). Diese neuere statistische Methode wertet in mehreren Schritten Daten über die Zusammensetzungen der Wählerschaft und der Stichprobe aus:

- Im ersten Schritt werden Idealtypen definiert (beispielsweise 18-35 jährige Bernerinnen, die in der Stadt wohnen).
- Die Daten der Stichprobe werden mit Mehrebenenmodellen analysiert, um zu berechnen, wie stark ein Idealtyp die Vorlage unterstützt.
- Im letzten Schritt werden die Vorhersagen pro Idealtyp gemäss der Zusammensetzung der Stimmbewölkerung gewichtet (Poststratifikation).

Wir haben in unseren Forschungsarbeiten diese neuere statistische Methode auf die Schweiz angewendet (siehe [hier](#) und [hier](#)). Die Auswertung der Tamedia-Abstimmungsumfragen ist in diesem Sinne ein Spin-off unserer Forschung. Auch wenn die konkrete methodische Durchführung anspruchsvoll ist, die Intuition der Methodik ist einfach: Das Augenmerk liegt im Vergleich zur Standardmethode nicht im Generieren der Umfrage, sondern in der Auswertung der grossen Onlinestichproben (> 10'000 Befragte in unserem Fall). Wir analysieren und gewichten die Daten der Onlinestichproben in einem mehrstufigen Verfahren, um eine möglichst präzise Aussage über die Unterstützung in der Stimmbewölkerung zu schätzen. Die gewichtete Analyse von Onlinedaten ist ebenfalls mit Herausforderungen und Fallstricken verbunden. Eine Schwäche unseres Verfahrens ist, dass wir nur sehr schwache Aussagen über die Stimmbeteiligung machen können. Ebenso ist die Durchführung des Gewichtungsverfahrens mit mehreren Ermessensentscheidungen verbunden. Wir überprüfen und hinterfragen unsere Entscheide permanent. Eine wichtige Annahme ist, dass sich beispielsweise die 18-35 jährigen Bernerinnen, die eine Berufslehre absolviert haben, in der Stadt leben, mit der SP sympathisieren und an der Tamedia Umfrage teilnehmen, nicht systematisch von den 18-35 jährigen Bernerinnen mit den selben Merkmalen unterscheiden, die nicht an der Umfrage teilnehmen. Bei jungen Leuten ist diese Annahme wohl unproblematisch, nicht aber bei älteren Wählerinnen.

Kontrovers diskutiert wird, inwieweit die offene Form der Teilnahme, wie wir sie in den Tamedia-Abstimmungsumfragen durchführen, ein zusätzliches Problem darstellt. YouGov und andere Institute führen Onlineumfragen ohne offene Teilnahmen durch. Eine wichtige Studie zu diesem Thema kommt ebenfalls von Andrew Gelman und Kollegen. In Zusammenarbeit mit Microsoft Research haben die Statistiker der Columbia University eine offene Umfrage über die Xbox-Konsole durchgeführt. Die Stichprobe von Xbox-Spielern



unterscheidet sich selbstverständlich stark von der Gesamtbevölkerung. Über die Anwendung fortgeschrittener Gewichtungsverfahren, konnten die Forscher jedoch präzise Aussagen über die Unterstützung im US Wahlkampf berechnen ([hier](#) ist der Blogeintrag dazu; [hier](#) das Papier). Trotzdem ist dieses Thema für uns von grosser Bedeutung. Wir schliessen jeweils in einem mehrstufigen Verfahren über 15% der Antworten aus, damit wir für die Analyse eine qualitativ hochstehende Stichprobe haben. Neben der Gewichtung in der Analyse ist die sorgfältige Generierung und Bereinigung der Daten ebenfalls von grosser Bedeutung.

Fazit: Mit Onlineumfragen können sehr grosse Stichproben effizient generiert werden. Komplexe Gewichtungsverfahren ermöglichen es, Aussagen über die Gesamtbevölkerung zu machen. Die Bereinigung der Daten und die Anwendung der mehrstufigen Gewichtungsverfahren sind aber mit mehreren Ermessensentscheidungen verbunden.



Bisherige Umfragen und Vergleich mit Endergebnis

Die Qualität von Umfragen zu bestimmen ist kein leichtes Unterfangen, da die Unterstützung von Vorlagen sich über die Zeit ändert. Die folgenden Daten hier erlauben jeder Leserin und jedem Leser sich ein eigenes Bild zu machen.

Es gibt zwei mögliche Ansätze, um Umfragen mit dem letztendlichen Abstimmungsergebnis zu vergleichen. Im ersten Ansatz wird die letzte Umfrage vor der Abstimmung mit dem Abstimmungsergebnis verglichen. Dieser Vergleich ist informativ unter der Annahme, dass sich die Meinungen in den letzten zwei bis drei Wochen zwischen der letzten Umfragewelle und der Abstimmung nicht mehr ändern. Eine weitere Möglichkeit besteht darin, die letzten beiden Umfragewellen linear zu extrapolieren. Dieser zweite Ansatz geht davon aus, dass der Meinungswandel konstant ist und sich auch im Zeitraum zwischen der letzten Umfragewelle und der Abstimmung fortsetzt. Beide Ansätze beruhen auf Annahmen. Ob die jeweilige Annahme in einem konkreten Fall zutrifft oder nicht ist eine empirische Frage. Im folgenden werden beide Gegenüberstellungen präsentiert.

Tabelle 1: Letzte Umfrage und tatsächliches Resultat

Vorlage	Tamedia Ja %	SRG Trendumfrage Ja %	Abstimmungsergebnis	Näher am Abstimmungsergebnis
Masseneinwanderung	49	43	50	Tamedia
FABI	46	56	62	SRG Trendumfrage
Abtreibung	28	36	30	Tamedia
Gripen	43	44	47	SRG Trendumfrage
Pädophilen Initiative	66	59	64	Tamedia
Mindestlohn	30	30	24	Tamedia/SRG Trendumfrage
Med. Grundversorgung	76	71	88	Tamedia
Öffentliche KK	48	38	38	SRG Trendumfrage
MWST Gastro	37	41	29	Tamedia
Pauschalbesteuerung	44	42	41	SRG Trendumfrage
Ecopop	36	39	26	Tamedia
Gold	28	38	23	Tamedia
Familieninitiative	39	40	25	Tamedia
Energieinitiative	12	19	8	Tamedia
Erbschaftsteuer	29	34	29	Tamedia
Präimplantationsdiagnostik	56	46	62	Tamedia
Radio & TV Gesetz	41	43	50	SRG Trendumfrage
Stipendieninitiative	39	38	28	SRG Trendumfrage
Durchsetzungsinitiative	43	46	41	Tamedia
Zweite Gotthardröhre	54	56	57	SRG Trendumfrage
Heiratsstrafe abschaffen	49	53	49	Tamedia
Skeplationsstopp	39	31	40	Tamedia
Asylgesetz	55	60	67	SRG Trendumfrage
Milchkuh	41	40	29	SRG Trendumfrage
Bedingungsloses Grundeinkommen	37	26	23	SRG Trendumfrage
Pro Service Public	48	40	32	SRG Trendumfrage
FPMG	55	keine Umfrage	62	×
Grüne Wirtschaft	42	51	32	Tamedia
AHVplus	49	41	41	SRG Trendumfrage
NDG	58	53	41	Tamedia
Atomausstieg	57	48	46	SRG Trendumfrage
Erleichterte Einbürgerung	55	66	60	Tamedia
Finanzierung Strassenverkehr	60	62	62	SRG Trendumfrage
USR III	45	45	41	Tamedia/SRG Trendumfrage

Wird die letzte Umfrage vor den Abstimmungen mit dem Endergebnis verglichen (siehe Tabelle 1), liegt die Tamedia Abstimmungsumfrage in mehr als der Hälfte (52%) aller Abstimmungen näher am amtlichen Endergebnis als die SRG Trendumfrage.



Ein weiterer Vergleich besteht darin, die letzten zwei Umfragewellen linear zu extrapolieren – dies wird normalerweise als Trend bezeichnet. In Tabelle 2 sind die letzten beiden Umfragewellen abgebildet und das erwartete Ergebnis basierend auf dem linearen Trend. Damit dieser Vergleich valide ist bedarf es der Annahme, dass sich die Unterstützung im Laufe einer Kampagne linear zur Zeit verändert. Die Resultate sind bei dieser Gegenüberstellung noch deutlicher – in drei von fünf Fällen liegt der Tamedia-Abstimmungstrend näher am amtlichen Ergebnis als die SRG Trendumfrage.

Tabelle 2: Lineare Projektion und tatsächliches Resultat

Vorlage	SRG Trendumfrage		Tamedia		Projektion		Näher am Abstimmungs- ergebnis
	≈ 7 W vor Abstimmung	≈ 3 W vor Abstimmung	≈ 5 W vor Abstimmung	≈ 2 W vor Abstimmung	SRG Trendumfrage	Tamedia	
Masseneinwanderung	37	43	×	49	47.5	×	×
FABI	56	56	×	46	56.0	×	×
Abtreibung	35	36	×	28	36.8	×	×
Gripen	42	44	36	43	45.5	47.7	Tamedia
Pädophile	74	59	72	66	47.8	62.0	Tamedia
Mindestlohn	40	30	35	30	22.5	26.7	SRG Trendumfrage
Med Grundversorgung	66	71	75	76	74.8	76.7	Tamedia
Oeffentliche KK	40	38	49	48	36.5	47.3	SRG Trendumfrage
MWST Restaurant	41	41	46	37	41.0	31.0	Tamedia
Pauschalbesteuerung	48	42	49	44	37.5	40.7	Tamedia
Ecopop	35	39	46	36	42.0	29.3	Tamedia
Gold	44	38	38	28	33.5	21.3	Tamedia
EnergieMWSt	29	19	13	12	11.5	11.3	Tamedia
Familien	52	40	57	39	31.0	27.0	Tamedia
Erbsteuer	38	34	27	29	31.0	30.3	Tamedia
PfD	40	46	37	56	50.5	68.7	Tamedia
RTVG	46	43	36	41	40.8	44.3	Tamedia
Stipendien	49	38	47	39	29.8	33.7	SRG Trendumfrage
Durchsetzung	51	46	51	43	42.25	37.7	SRG Trendumfrage
Gotthardtunnel	64	56	56	54	50.0	52.7	Tamedia
CVP Heiratsinitiative	67	53	61	49	42.5	41.0	SRG Trendumfrage
Spekulationsstopp JUSO	48	31	40	39	18.3	38.3	Tamedia
AsylG	59	60	49	55	60.8	59.0	SRG Trendumfrage
Milchkuh	42	40	51	41	38.5	34.3	Tamedia
bdgsl GrundEink	24	26	33	37	27.5	39.7	SRG Trendumfrage
Pro Service Public	58	46	59	48	37.0	40.7	SRG Trendumfrage
FPMG	×	×	48	55	×	59.7	×
Grüne Wirtschaft	61	51	48	42	43.5	38.0	Tamedia
AHVplus	49	40	54	49	33.3	45.7	Tamedia
NDG	58	53	59	58	49.3	57.3	Tamedia
AtomAusstieg	57	48	56	57	41.3	57.7	SRG Trendumfrage
Erleichterte Einbürgerung	74	66	51	55	60.0	57.7	SRG Trendumfrage
Finanzierung Strassenverkehr	60	62	52	60	63.5	65.3	SRG Trendumfrage
USR III	50	45	44	45	41.3	45.7	SRG Trendumfrage

Beide Vergleiche – basierend auf unterschiedlichen Annahmen – weisen in dieselbe Richtung: Die Tamedia Abstimmungsumfragen sind im Vergleich sehr gut und schneiden besser ab als alternative Umfragen - können im Einzelfall aber wie jede Umfrage vom Ergebnis abweichen. Diese Gegenüberstellungen zeigen die Leistungsfähigkeit des in den Tamedia-Umfragen verwendeten innovativen modellbasierten Ansatzes.



Oft gefragt (Frequently Asked Questions)

Uns werden immer wieder Fragen zu unserem Vorgehen, den Resultaten und unserer Zusammenarbeit mit Tamedia gestellt. Wir versuchen hier die wichtigsten Fragen – so weit wie möglich – zu beantworten. Sie können uns gerne weitere Fragen stellen, die wir in überarbeiteten Versionen dieses Dokumentes ebenfalls beantworten werden (sofern diese von allgemeinem Interesse sind).

Wie präzise sind unsere Umfrageergebnisse?

Diese Frage ist schwierig zu beantworten, weil die "wahren" Werte unbekannt sind. Der Vergleich der letzten Umfragewerte, die jeweils 10 Tage vor der Abstimmung publiziert werden, mit dem tatsächlichen Abstimmungsergebnis ist ein guter Plausibilitätstest und zeigt, dass unsere Analysen verhältnismässig präzise sind. Bei über der Hälfte der Vorlagen war die letzte Tamedia-Abstimmungsumfrage näher am Endresultat als die letzte SRG-Trendumfrage, welche beide zehn Tage vor dem Abstimmungstermin publiziert werden (siehe Tabelle 1).

Wir analysieren unsere Resultate und Methodik permanent und integrieren die Erkenntnisse der Literatur systematisch in unsere Arbeit. Wir möchten aber Journalistinnen, Politikern und Bürgerinnen nahe legen, Umfrageresultate generell vorsichtig zu interpretieren. Umfrageergebnisse sind – je nach Qualität der Umfrage und Analyse – bessere oder weniger gute Indikatoren. Zudem sind Aussagen zur Unterstützung innerhalb von Unterkategorien (wie z.B. Kantone und Parteien) weniger genau als Schätzungen zur Gesamtbevölkerung und sollten daher besonders zurückhaltend interpretiert werden.

Wie funktioniert unsere Zusammenarbeit mit Tamedia?

In der Zusammenarbeit mit Tamedia beschränken wir uns auf die Durchführung der Umfrage und die Analyse der Daten. Die journalistische Aufarbeitung der Umfrageresultate liegt in der Verantwortung der JournalistInnen der Tamedia Titel. Wir haben bewusst den Entscheid gefällt, Abstimmungskampagnen nicht zu kommentieren und unsere Ergebnisse öffentlich nicht mit politischen Ereignissen zu plausibilisieren. Wir beschränken uns auf unsere Kernkompetenz, die methodisch saubere Durchführung der Analysen.

Wie genau gewichten wir?

Grundsätzlich geht es darum die Stichprobe der Struktur der Stimmbevölkerung anzupassen. Wir berücksichtigen politische, geographische und demographische Informationen. Natürlich ist der "heilige Gral" gewichteter Analysen die genaue Durchführung der mehrstufigen Analyse. Die Analysen werden von uns ständig überprüft und hinterfragt. Es gibt nicht *die* richtige Lösung – sondern bessere und weniger gute. Das genaue Gewichtungsschema ist vertraulich.

Kann die Tamedia-Abstimmungsumfrage manipuliert werden?



Da die Umfrage offen ist kommt immer wieder die Frage auf, ob unsere Umfrage manipuliert werden kann. Es gibt zwei Arten von Manipulationen. Einerseits versuchen Individuen die Umfrage zu manipulieren, indem sie mehrfach teilnehmen und inkorrekte Angaben machen; andererseits sind wir mit technisch versierteren Manipulationen von organisierten politischen Akteuren mit genügend Ressourcen konfrontiert. Ersteres ist unproblematisch. Für die Bekämpfung der zweiten Art der Manipulation betreiben wir einen erheblichen Aufwand. Wir haben zusammen mit unserem IT-Mitarbeiter eine Reihe von Sicherheitselementen für die Qualitätssicherung der Datenerhebung eingebaut. In zahlreichen Schritten überprüfen wir die Daten. Dies führt dazu, dass wir mehr als 15% der Antworten nicht berücksichtigen, da eine oder mehrere Warnelemente positiv sind. Aus nachvollziehbaren Gründen können wir die einzelnen Sicherheitselemente nicht offenlegen. Grundsätzlich gilt: Es ist nicht unmöglich die Umfrage zu manipulieren, aber die Kosten und Kompetenzen, die dafür notwendig wären, sind beträchtlich und wir arbeiten kontinuierlich daran, es noch schwieriger und umständlicher zu machen.

Beeinflussen Umfragen die Abstimmungsergebnisse?

Als einer von zahlreichen Faktoren beeinflussen Umfrageergebnisse die Abstimmungskampagnen. Aber beeinflussen sie auch die Abstimmungsergebnisse? Theoretisch sind viele Effekte denkbar – insbesondere auch Einflüsse, die sich gegenseitig aufheben. Ein Umfrageergebnis kann sowohl die vermeintlichen Abstimmungsgewinner als auch die Verlierer mobilisieren. Aus demokratietheoretischer Sicht, wird Mobilisierung und damit einhergehende hohe Stimmbeteiligung generell positiv bewertet. Problematisch könnte sein, wenn Umfrageergebnisse wesentlich stärker das eine Lager (de)mobilisieren. Mehrere Studien kommen jedoch zum Schluss, dass Umfragen höchstens einen vernachlässigbar kleinen Effekt auf die Abstimmungsergebnisse haben (siehe [hier](#) und [hier](#)).

Ist es sinnvoll Umfrageresultate zu publizieren?

Für viele Bürgerinnen ist die Meinung im Stimmvolk eine wichtige Information in der Entscheidungsfindung (neben den Argumenten und Meinungen von Freundinnen, Arbeitskollegen, Parteien, Parlament, Regierung, Interessenorganisationen etc.). Die Stimmungslage im Volk erfahren sie durch öffentlich publizierte Umfragen. Noch wichtiger in diesem Zusammenhang ist, dass wir nicht nur Fragen nach dem Stimmverhalten fragen, sondern auch nach den überzeugendsten Argumenten. Gerade die Auswertungen zu den Argumenten werden in der medialen Berichterstattung intensiv diskutiert und helfen den Stimmbürgerinnen sich über die Vor- und Nachteile einer Vorlage zu informieren. Wir möchten ebenfalls betonen, dass Umfragen von professionellen Akteuren in den Abstimmungskampagnen in Auftrag gegeben werden. Die Resultate dieser Umfragen werden gar nicht publiziert (oder nur selektiv mit der Absicht die Abstimmung in eine bestimmte Richtung zu beeinflussen). Wer sich also gegen öffentliche Meinungsumfragen stellt, unterstützt, dass nur finanzstarke Politprofis, die in Kampagnen für die eine oder andere Seite engagiert sind, Umfrageanalysen über die Meinungsbildung im Volk zur Verfü-



gung haben. Wir bevorzugen eine öffentliche und somit allen zugängliche Diskussion von politischen Umfrageanalysen.

Fazit

Gerade weil wir selber Umfragen durchführen und uns intensiv mit der Praxis und Literatur beschäftigen, empfehlen wir Umfragen vorsichtig zu interpretieren. Über die Jahre haben sich die Problem mit der Standardmethode (Telefonumfragen) vergrößert, während wissenschaftliche Fortschritte und das veränderte Kommunikationsverhalten Alternativen generiert haben. Die Erfahrungen in der Schweiz und den USA zeigen, dass Onlineumfragen nicht kritischer zu betrachten sind als Telefonumfragen. Dieses Dokument hat die statistischen Grundlagen der noch jungen mehrstufigen Gewichtungsanalysen von Onlinedaten diskutiert und oft gestellte Fragen bezüglich unserer Analyse von Tamedia-Abstimmungsumfrage beantwortet. Wir hoffen mit diesen Ausführungen zu einem kritischen und informierten Austausch über Abstimmungsumfragen beizutragen.



Über die Autoren

Lucas Leemann ist Associate Professor (Reader) an der Universität Essex und Lehrbeauftragter an der Universität Zürich. Er hat seine Doktorarbeit an der Columbia Universität geschrieben und unterrichtet komplexe statistische Anwendungen an verschiedenen Universitäten (University College London, Essex, Zürich, Waseda Tokio).

Fabio Wasserfallen ist Assistenzprofessor an der Universität Salzburg und Gastprofessor an der Universität Zürich. Er war für einjährige Forschungsaufenthalte in Harvard und Princeton und ko-kordiniert das Horizon 2020 Projekt "EMU Choices", welches die politischen Zusammenhänge der Eurokrise untersucht.