

Working Paper Energie und Ressourcen

Felduntersuchung zu Behavioral Energy Efficiency Potentialen von privaten Haushalten

Autoren: *Adrian Beyertt*¹, *Paul Verwiebe*², *Stephan Seim*², *Filip Milojkovic*¹,
*Joachim Müller-Kirchenbauer*²

¹ Fresh Energy GmbH

² Fachgebiet Energie- und Ressourcenmanagement, TU Berlin

Key-Words: <ul style="list-style-type: none">– Behavioral Energy Efficiency– Smart Meter– NILM– Disaggregation– Feedback– Residential Electricity Consumption	Abstract: In dieser Untersuchung wird das Potential von Behavioral Energy Efficiency (BEE) in privaten Haushalten analysiert. Dazu wurde ein Feldexperiment von Oktober 2019 bis Januar 2020 durchgeführt, bei dem 100 Haushalte einen monatlichen Energiereport erhalten haben, der auf ihren Smart Meter Daten (0,5 Hz) und deren Disaggregation basierte. Der Report enthielt zusätzliche Informationen zu jenen, die bereits über eine App zur Verfügung gestellt wurden, wie z. B. normative und historische Vergleiche. Im Ergebnis konnte keine signifikante Änderung im Stromverhaltensverhalten festgestellt werden. Allerdings verwendeten über 75% den Report regelmäßig und 60% der Teilnehmer gaben an, dass die Erkenntnisse zu einer Änderung ihres Stromverhaltensverhalten geführt haben. Weiterführende Forschung sollte sich mit der Untersuchung von kurzfristigen Effekten von Feedback oder der Weiterentwicklung von einzelnen Feedbackkomponenten befassen.
---	--

Inhalt

1	Einleitung.....	2
2	Theoretische Grundlagen.....	2
2.1	Behavioral Energy Efficiency	2
2.2	Smart Meter und Disaggregation von Stromverbrauchsdaten (NILM)	3
3	Experiment	4
3.1	Design.....	4
3.2	Ergebnisse.....	8
4	Diskussion der Ergebnisse	10
5	Fazit und Ausblick.....	11
A. Anhang:	Statistischer Überblick der Kontrollvariablen.....	15
B. Anhang:	Monatlicher Report	16
C. Anhang:	Fließschema der abschließenden Online-Umfrage	17
D. Anhang:	Statistische Aufbereitung der Gesamtverbrauchsdaten	18

1 Einleitung

Die menschliche Entscheidungsfindung und folglich die Erklärung des individuellen Verhaltens ist komplex. In Bezug auf den Energieverbrauch ist es ohne hochaufgelöste Visualisierung und individuelle Interpretation der Daten schwierig, die Auswirkungen des eigenen Handelns zu begreifen. Die Digitalisierung der Erfassung und Übermittlung von Stromverbrauchsdaten im Sekundentakt mit Smart Metern nahezu ohne Zeitverzögerung bietet die Chance, die Datenanalyse sowie die Interaktion mit dem Endkunden auf Nachhaltigkeit und Energieeffizienz auszurichten.

Der Markt für digitale Messtechnik mit Kommunikationstechnologie ist in Deutschland noch im Aufbau und auch Studien mit realen Kunden gibt es bislang kaum (ein Untersuchungsbeispiel aus Deutschland ist etwa [1]). Erste Ergebnisse, bspw. aus den USA, zeigen, dass insbesondere kurzfristiges Feedback ein wirksames Konzept ist [2]. Die gängige These lautet: Werden die Stromverbrauchsdaten entsprechend verhaltensökonomischer Konzepte aufbereitet und privaten Konsumenten zur Verfügung gestellt, können diese ihr Verhalten energieeffizienter gestalten [3, 4].

Die vorliegende Arbeit dokumentiert die Entwicklung eines Feldexperiments mit datenbasiertem Feedback in Kooperation mit der Fresh Energy GmbH im Zeitraum Oktober 2019 bis Januar 2020. Das Produkt von Fresh Energy ist eine App, in der Kunden ihren Stromverbrauch in (nahezu) Echtzeit einsehen können. Algorithmische Disaggregation der mit 0,5 Herz erfassten Stromlast am Haushaltszähler, das sogenannte "Non-Intrusive Load Monitoring" (NILM), ermöglicht die Berechnung und Darstellung des Verbrauchs einzelner Geräte ohne separate Messung.

Auf Basis von monatlichen Verbrauchsdaten des Haushalts und der einzelnen (algorithmisch erkannten) Geräte wurde in einem iterativen Vorgehen mit mehreren Umfragen und Kundeninterviews ein Report mit zusätzlichen Kontextinformationen zum Stromverbrauch erstellt, der im Rahmen des Experiments 100 Haushalten (der Testgruppe) per E-Mail zur Verfügung gestellt wurde. Die Zielstellung des Experiments ist die Analyse des Stimulus in Form des monatlichen Reports auf den Stromverbrauch privater Haushalte. Die Hypothese ist also, dass der Stromverbrauch der Testgruppe im Vergleich zur Kontrollgruppe sinkt. Die für dieses Paper analysierten Daten und die gewählte Methodik lassen keine Bestätigung der Hypothese zu. Es kann nicht festgestellt werden, dass die Testgruppe gegenüber der Kontrollgruppe Strom einspart. Allerdings weisen qualitative Ergebnisse aus dem Entwicklungsprozess und dem Feedback der Testgruppe darauf hin, dass der Report als hilfreich für eine Änderung des Stromverhaltens einzustufen ist.

Das folgende Kapitel stellt den theoretischen Hintergrund zu Behavioral Energy Efficiency dar und definiert die technische Ausstattung der Haushalte sowie den Ansatz der algorithmischen Disaggregation, das sogenannte NILM. Kapitel 3 beschreibt das durchgeführte Experiment und den Datensatz mit Ergebnissen der Analyse. Darauf folgen eine Diskussion der Ergebnisse in Kapitel 4 und ein Fazit in Kapitel 5.

2 Theoretische Grundlagen

2.1 Behavioral Energy Efficiency

Entgegen der neoklassischen Ökonomie beschreibt die Verhaltensökonomie („Behavioral Economics“) das Entscheidungsverhalten des Menschen vor dem Hintergrund unvollständiger Informationen und eingeschränkter Rationalität. Werden Konzepte aus der Verhaltensökonomie auf den Energiebereich übertragen, wird von Behavioral Energy Efficiency (BEE) gesprochen.

Einer der zentralen Untersuchungsgegenstände der Behavioral Energy Efficiency ist die Frage, wie Stromverbrauchsdaten aufbereitet werden können, um ein energieeffizienteres Verhalten zu unterstützen. Das Individuum stößt dabei auf verhaltensbezogene Barrieren, die ein energieeffizienteres

Verhalten erschweren. Zum einen haben Menschen atypische Präferenzen. Diese äußern sich beispielsweise im Ausstattungseffekt (“endowment effect”) [5, 6], in einer Tendenz zur Gegenwart („present bias“) [7–9] und der Motivation eines Individuums [10–13]. Zum anderen ist das Entscheidungsverhalten bedingt rational. Hier sind Heuristiken zur Entscheidungsfindung [14, 15], Trägheit [16–18] und das Framing von Informationen [19, 20] wichtige Konzepte aus der Verhaltensökonomik.

Im Zusammenhang mit Smart Meter Daten sticht eine Methodik hervor, um diese verhaltensbezogenen Barrieren zu überwinden: Feedback. Unter Beachtung von verhaltensökonomischen Ansätzen können Energieverbrauchsdaten aufbereitet werden und einen Beitrag dazu leisten, das Verhalten von Menschen energieeffizienter zu machen. Eine Unterteilung von Feedbackarten erfolgte in einer Meta-Analyse durch Fischer [21], welche Experimente basierend auf Stromverbrauchsdaten untersuchte. Feedback kann demnach durch folgende Komponenten charakterisiert werden:

- Frequenz und Dauer
- Inhaltliche Ausgestaltung
- Aufschlüsselung auf einzelne Geräte
- Gewähltes Kommunikationsmedium
- Verwendung von Vergleichen
- Zusätzliche Informationen und Maßnahmen

In der Meta-Analyse zu Stromverbrauchsdaten und Feedback stellten Carrie Armel et al. [4] fest, dass die Stromeinsparungen je nach Feedbackzusammensetzung mehr als 12% betragen können. Ein oft zitiertes Beispiel zum Thema Feedback ist der Home Energy Report von Opower [2, 22], welcher u.a. historische und normative Vergleiche verwendet. Ein Beispiel aus dem deutschsprachigen Raum ist die Untersuchung von Schleich et al. [1], welche die Verbrauchsdaten u.a. in einem Webportal zugänglich machte. Hier konnten Stromeinsparungen von etwa 3,7% erzielt werden.

2.2 Smart Meter und Disaggregation von Stromverbrauchsdaten (NILM)

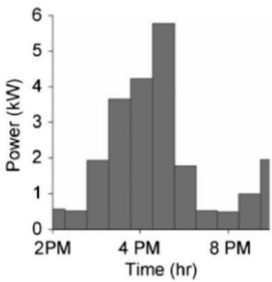
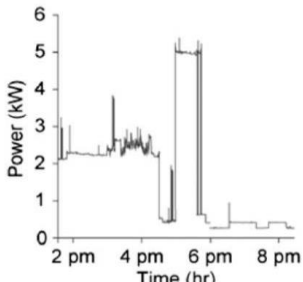
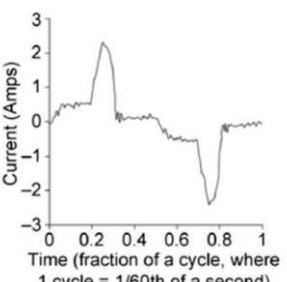
Im Allgemeinen wird in Deutschland bei der Modernisierung der Stromzählertechnologie vom „Smart Meter Rollout“ gesprochen, der im Messstellenbetriebsgesetz (MsbG) geregelt ist. So wird im MsbG zwischen einem intelligenten Messsystem (iMSys) und einer modernen Messeinrichtung (mMe) unterschieden. Eine mMe ist lediglich ein elektronischer Zähler mit einem Display, welcher die Verbrauchswerte direkt am Gerät anzeigen kann. Allerdings ist diese nicht fähig, die Daten zu übertragen. Ein iMSys besteht aus einer mMe und einem Smart-Meter-Gateway, welches die Daten an weitere Marktteilnehmer senden kann. Standardmäßig sollen aktuell mindestens Verbrauchsdaten im Viertelstundenabstand erfasst und kommuniziert werden. Da bei Haushalten mit einem Jahresverbrauch von unter 6.000 kWh bis auf weiteres keine Pflicht zum Einbau eines iMSys besteht, ist ein flächendeckender Rollout nur mit für alle Marktteilnehmer relevanten datenbasierten Mehrwertdiensten auf freiwilliger Basis zu erwarten. Beispiel für solch einen Mehrwert ist die automatisierte Erkennung des Verbrauchs einzelner Geräte ohne separate Messung.

Das Herunterbrechen einer Lastkurve auf einzelne Verbraucher (“Disaggregation”) wird durch sogenannte NILM-Algorithmen (“Non-Intrusive-Load-Monitoring”) realisiert. Aus verhaltensökonomischer Sicht haben gerätespezifische Verbrauchsinformationen in Verbindung mit kurzfristiger Visualisierung das Potential, das Handeln stärker mit dessen Auswirkung zu verknüpfen. Ein erstes NILM-Verfahren wurde bereits Anfang der 90er Jahre von Hart [23] vorgestellt. „Non-Intrusive“ bezieht sich dabei auf die Messung einer Gesamtlast mit verschiedenen Verbrauchern (z.B. ein privater Haushalt) und steht damit im Gegensatz zu intrusiven Verfahren, die jedes Gerät einzeln messen („Intrusive-Load-Monitoring“).

NILM-Ansätze untersuchen die Gesamtlastkurve auf Signaturen einzelner Geräte mit Algorithmen aus dem Bereich des maschinellen Lernens. Durch einen Abgleich mit einer Datenbasis, die als initiales Training für die Algorithmen dient, können die Signaturen einem elektrischen Verbraucher zugewiesen werden. Größtenteils wird dabei auf Verfahren aus der Mustererkennung zurückgegriffen, wobei vor allem Ansätze basierend auf dem Hidden-Markov-Model [24, 25] und neuronale Netzwerke [26] zum Einsatz kommen.

Wichtig sind dabei die Veränderungen in den Messdaten über den Zeitverlauf. Diese spiegeln die Zustandsänderungen der elektrischen Verbraucher im System wider. Je höher dabei die verwendete Messfrequenz ist, desto genauer können einzelne Verbraucher identifiziert werden. Eine Übersicht der Implikationen der Datenfrequenz für NILM-Verfahren bietet die Meta-Analyse von Carrie Armel et al. [4], in der etwa 40 verschiedene Studien analysiert wurden (vgl. Tabelle 1).

Tabelle 1: Übersicht der Auswirkung der gewählten Datenfrequenz auf die Charakteristika der NILM-Algorithmen und die Erkennung von Geräten (abgewandelte Darstellung, angelehnt an Carrie Armel et al.) [4, S. 221]

Datenfrequenz	1h-15min	1min-1s (1Hz)	> 1Hz
Graphisch			
Features der Algorithmen	Optisch erkennbare Muster; Dauer und Uhrzeit der Verwendung	Stationärere Zustände und Übergänge zwischen einzelnen Lastzuständen	Spannung und Stromstärke, komplexere Wellenformen
Erkennung von Geräten	Drei Kategorien: Lasten mit Korrelation zur Außentemperatur, Grundlast und zeitabhängige Lasten	Bis zu zehn Gerätetypen: Kühlschrank, Geschirrspüler, Waschmaschine, Trockner, etc.	Bis zu 100 verschiedene Geräte sowie Unterscheidung zwischen zwei einzelnen Geräten einer Kategorie

Vorteile von NILM gegenüber intrusiven Verfahren sind geringere Kosten und eine bessere Skalierbarkeit, da nur ein einziger Sensor benötigt wird. Die von Fresh Energy eingesetzte NILM-Technologie beruht auf Verbrauchsdaten mit einer Frequenz von 0,5 Hz (ein Wert alle 2 Sekunden) für die Disaggregation auf einzelne Geräte. NILM-Verfahren bieten allerdings keine absolute Zuverlässigkeit und Genauigkeit, also die Übereinstimmung der Interpretation des Algorithmus mit der Realität, die je nach Verfahren und Geräteklasse zwischen 75% und 95% liegen [27].

3 Experiment

3.1 Design

Die Zielstellung des Experimentes ist es, herauszufinden, ob ein Stimulus in Form eines monatlichen Reports den Stromverbrauch privater Haushalte senken kann. Die Hypothese ist also, dass der Stromverbrauch einer Testgruppe im Vergleich zur Kontrollgruppe sinkt. In einem ersten Schritt wird das Design des Experimentes näher erläutert.

Definition der Variablen

Grundsätzlich sind drei Arten von Variablen bei einem Experiment zu beachten: unabhängige und abhängige Variablen sowie Kontrollvariablen. Eine unabhängige Variable ist ein Merkmal, dessen Effekt

in Abhängigkeit seiner Ausprägungen auf die abhängige Variable untersucht werden soll. Kontrollvariablen sind Merkmale, die durch den Datensatz selbst gegeben und für den Zusammenhang zwischen abhängiger und unabhängiger Variable wichtig sind.

Für die vorliegende Untersuchung ist demnach der monatliche Stromverbrauch die abhängige Variable, welche monatlich aggregiert wird. Die Verbrauchsdaten sind kardinal skaliert und werden in kWh erfasst. Der Versand des monatlichen Reports wird als Stimulus interpretiert und geht somit als unabhängige Variable in das Modell ein, dessen Einfluss auf den Verbrauch gemessen wird.

Darüber hinaus wird für die Testgruppe das Engagement während des Experiments als Einflussfaktor berücksichtigt. Das Engagement der Kunden wird aus der Interaktion mit dem Report abgeleitet. Durch ein einfaches Scoring-System soll gemessen werden, ob die E-Mail ignoriert wurde (0 Punkte) oder ob die E-Mail mit dem Report geöffnet wurde (1 Punkt). Am Ende des Experimentes wird ein Durchschnittswert gebildet, der das durchschnittliche Engagement des Teilnehmers während des Experiments darstellt.

Stichprobenumfang

Die zur Verfügung stehende Grundgesamtheit ergab sich aus den 1.616 Beta-Testern von Fresh Energy. Von diesen gaben 326 Personen ihr Einverständnis für eine Teilnahme. Aus diesen wurde randomisiert eine Testgruppe von 100 Haushalten erstellt, welche den Report für die Dauer des Experiments erhielten. Während des Experiments traten bei 12 Haushalten der Testgruppe Soft- oder Hardware-Probleme auf, sodass letztlich 88 vollständige Datensätze für die Testgruppe zur Verfügung standen. Diese 88 Haushalte bilden die TG1 und werden weiter durch die Variable Engagement unterteilt. Während des Experimentes öffneten 53 Haushalte die E-Mail mit dem Report jeden Monat. Diese formen die Gruppe TG2. Die restlichen 35 Haushalte mit geringerem Engagement bilden die TG3.

Die Kontrollgruppe wurde aus den restlichen 1.516 Datensätzen initialisiert. Für die Auswertung wurden hier insgesamt 479 Haushalte aussortiert. Gründe dafür waren eine eigene Stromproduktion (Producer), unvollständige Datensätze sowie Hard- und Softwareprobleme. Einen Überblick über die finale Zusammensetzung der zur Auswertung verwendeten Gruppen gibt Abbildung 1.

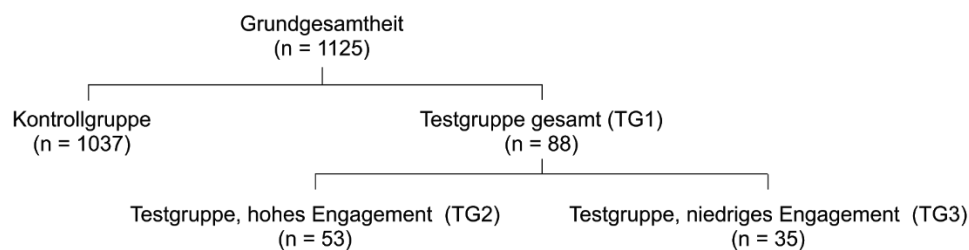


Abbildung 1: Überblick der gebildeten Gruppen im Experiment

Als Kontrollvariablen standen das Alter, das Geschlecht und die Postleitzahl zur Verfügung. Allerdings beziehen sich Angaben zu Alter und Geschlecht lediglich auf die Person, die den Vertrag bei Fresh Energy abgeschlossen hat. Ein Rückschluss auf weitere Personen innerhalb des Haushalts ist nicht möglich. Eine statistische Aufbereitung der Kontrollvariablen ist in Anhang A angefügt.

Definition des Stimulus

Für das Experiment erhalten die Testkunden zusätzlich zur Fresh Energy App einen monatlichen Report. Dieser basiert auf den hochfrequenten Smart Meter Daten und disaggregierten Geräteverbräuchen. Er enthält historische und normative Vergleiche, Referenzen zum CO₂-Fußabdruck und konkrete Stromspartipps. Der Report wurde in einem mehrstufigen Prozess mit Kundenumfragen und -interviews entwickelt und kontinuierlich verbessert.

Abbildung 2 zeigt einen Ausschnitt aus dem Report. Die vollständige Version ist in Anhang B angefügt.

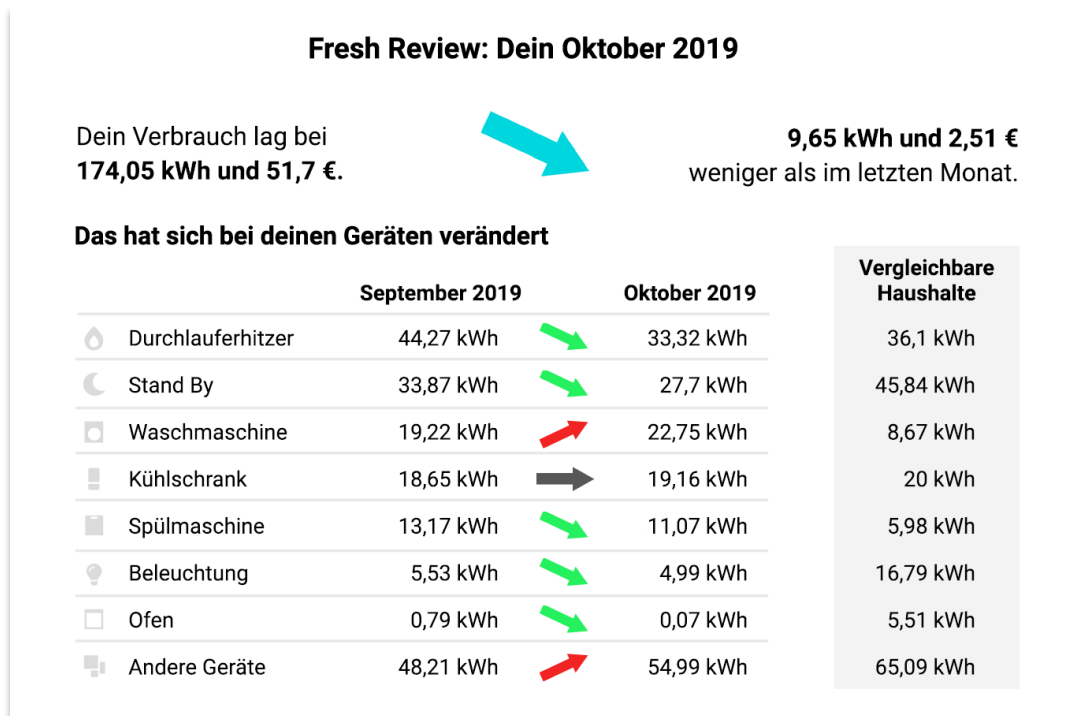


Abbildung 2: Ausschnitt aus dem monatlichen Report (der vollständige Report ist in Anhang B angefügt)

Zeitlicher Rahmen

Die Interaktion mit der Testgruppe war eng mit der Gestaltung des monatlichen Reports verbunden. Aus zeitlicher Perspektive waren die Intervalle des Stimulus per se gegeben, sodass einmal im Monat eine Kommunikation mit der Testgruppe erfolgte. Der Report wurde als PDF-Datei im Anhang einer E-Mail an die Testgruppe geschickt. Abbildung 3 gibt einen Überblick über den zeitlichen Ablauf des Experiments.

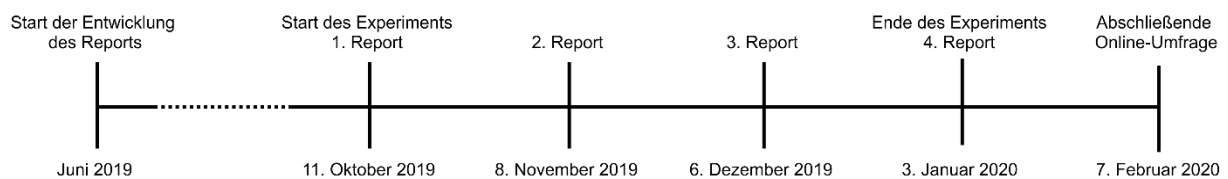


Abbildung 3: Zeitlicher Überblick über das Experiment

Planung der statistischen Auswertung: Modellbildung

Für das vorliegende Experiment ist ein Bottom-Up-Ansatz geeignet, der die monatlichen Verbrauchswerte in Relation zu einer Baseline setzt. Der gewählte Ansatz lehnt sich an die vom BMU entwickelten Methodiken zur Berechnung von Energieeinsparungen an [28]. Die Baseline simuliert, wie die Verbrauchsdaten sich entwickelt hätten, wenn keine Intervention stattgefunden hätte. Für die Laufzeit des Experiments bietet sich der durchschnittliche Tagesverbrauch der Kontrollgruppe als Baseline an. Es wird vereinfachend angenommen, dass die Kontroll- und Testgruppe im gleichen Maße durch externe und saisonale Effekte beeinflusst werden. Diese Vorgehensweise hat den Vorteil, die Komplexität zu reduzieren und die Anforderungen an die Menge der Messdaten zu senken. In Anlehnung an die Bottom-Up Methodik des BMU dient also der Verbrauchstrend der Kontrollgruppe als Baseline, um

eine hypothetische Verbrauchsentwicklung der Testgruppe ohne den monatlichen Report zu entwickeln. Abbildung 4 verdeutlicht die gewählte Vorgehensweise und Berechnungssystematik des Treatment-Effekts, also der Differenz zwischen der Verbrauchsentwicklung der Test- und Kontrollgruppe als Effekt des monatlichen Reports.

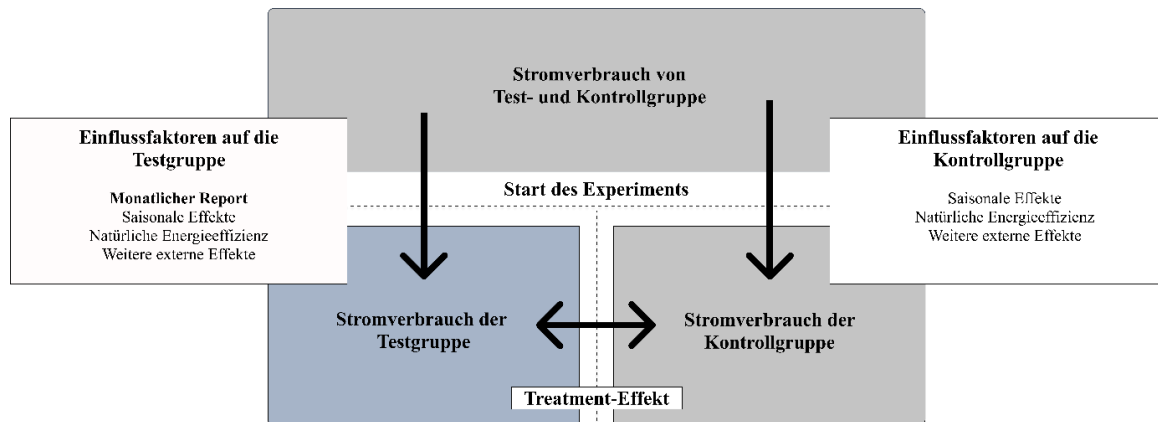


Abbildung 4: Gewählte Methodik der Berechnung des Treatment-Effekts

Entsprechend der Forschungshypothese ist bzgl. des Treatment-Effekts das Folgende zu untersuchen:

- H0: Der monatliche Report senkt den Stromverbrauch der Testgruppe nicht
- H1: Der monatliche Report senkt den Stromverbrauch der Testgruppe

Die Verbrauchswerte werden ausgewertet, indem zunächst der monatliche, durchschnittliche Tagesverbrauch berechnet wird. Dies berücksichtigt die Schwankungen in der Tagesanzahl zwischen den Monaten. Die Verbrauchswerte der Testgruppe werden sodann in Relation zu den Verbrauchswerten der Kontrollgruppe gesetzt. Die prozentuale Abweichung und deren zeitlicher Verlauf bestimmt den Treatment-Effekt. Diese Auswertung geschieht sowohl für den Gesamtverbrauch als auch für die disaggregierten Verbrauchsdaten.

Während des Experimentes wurde darüber hinaus überprüft, ob die Testkunden die E-Mail mit dem Report überhaupt öffnen. Hierdurch lässt sich qualitativ ableiten, inwiefern die Testgruppe die Reports verwendete. Im Anschluss an das Experiment folgte eine abschließende Kundenumfrage, welche am 6. Februar 2020 an die Kunden geschickt wurde. Einen Überblick über den Aufbau dieser Umfrage bietet Anhang C. Das Feedback der Teilnehmer wird dabei kurz qualitativ analysiert und vorgestellt.

Gütekriterien des Experiments

Die Generalisierbarkeit der vorliegenden Felduntersuchung, basierend auf Realdaten bestehender Haushaltskunden, ist eingeschränkt durch die Selbstauswahl der Teilnehmer ("self-selection bias"). Kunden von Fresh Energy sind bereits für ihren Stromverbrauch sensibilisiert. Auch das explizite Einverständnis zum Testen von neuen Features spricht für eine zusätzliche intrinsische Motivation der Teilnehmer. Außerdem ergibt sich durch die geringe Kontrollmöglichkeit von Störfaktoren eine geringere interne Validität als Nachteil, also die eindeutige Interpretierbarkeit der Ergebnisse. Ferner ist das Experiment nicht-reaktiv. Die Personen wissen also nicht, dass sie Teil eines Experimentes sind. Aus Sicht der Testkunden des Experimentes wurde lediglich kommuniziert, dass die Teilnehmer bei der Weiterentwicklung der Fresh Energy App neue Features testen dürfen.

Weiter charakterisieren sich Experimente durch ihre experimentelle oder quasi-experimentelle Untersuchungsart. Diese Unterscheidung klassifiziert die Auswirkung der personenbedingten Kontrollvariablen auf die Ergebnisse. Während die quasi-experimentelle Untersuchung natürlich ausgeprägte

Gruppen berücksichtigt, repräsentiert eine experimentelle Untersuchung eine ähnliche Verteilung der Kontrollvariablen innerhalb der Ausprägung unabhängiger Variablen. Das bedeutet für die vorliegende Arbeit, dass sich Test- und Kontrollgruppe in den Kontrollvariablen möglichst ähnlich sind. Aufgrund beschränkter Informationen zu den Gruppen kann dies allerdings nicht gewährleistet werden. Durch die randomisierte Auswahl der Testgruppe wird jedoch die Ähnlichkeit zur Kontrollgruppe angenähert. Die Untersuchung ist somit als experimentell einzustufen, was die interne Validität erhöht.

3.2 Ergebnisse

3.2.1 Analyse der Gesamtverbrauchsdaten

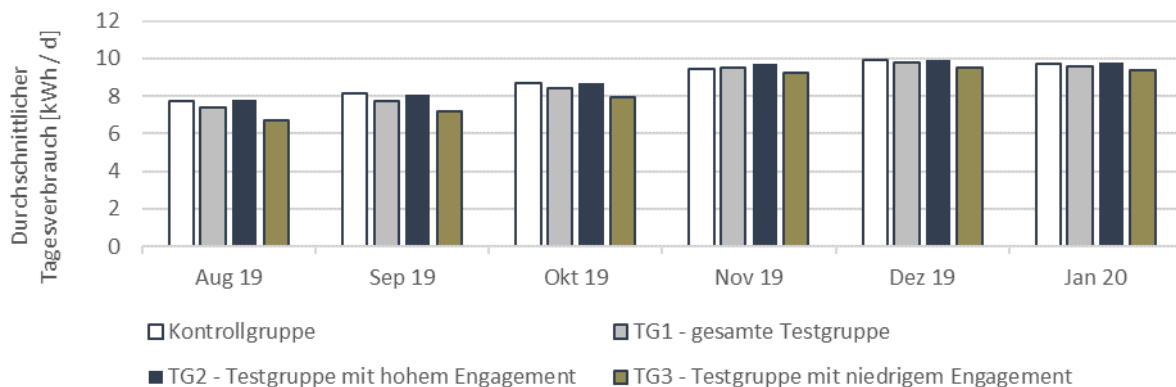


Abbildung 5: Vergleich des durchschnittlichen Gesamtverbrauchs von Test- und Kontrollgruppe im Zeitraum August 2019 bis Januar 2020

Abbildung 5 zeigt die Entwicklung des Gesamtverbrauchs im Zeitraum August 2019 bis Januar 2020. Es werden für jeden Monat die Verbrauchsdaten der Kontrollgruppe, der Testgruppe insgesamt, der Testgruppe „niedriges Engagement“ und der Testgruppe „hohes Engagement“ dargestellt. Eine weitere Perspektive auf die Gesamtverbrauchsdaten bietet Abbildung 6. Hier werden die Verbrauchsdaten der Testgruppe in Relation zur Kontrollgruppe gesetzt. Bis zum September 2019 war der Verbrauch der Testgruppe über 5% geringer als der Verbrauch der Kontrollgruppe. Mit dem Beginn des Experiments sinkt die Differenz und im November hat die Testgruppe insgesamt 0,85% mehr Strom als die Kontrollgruppe verbraucht.

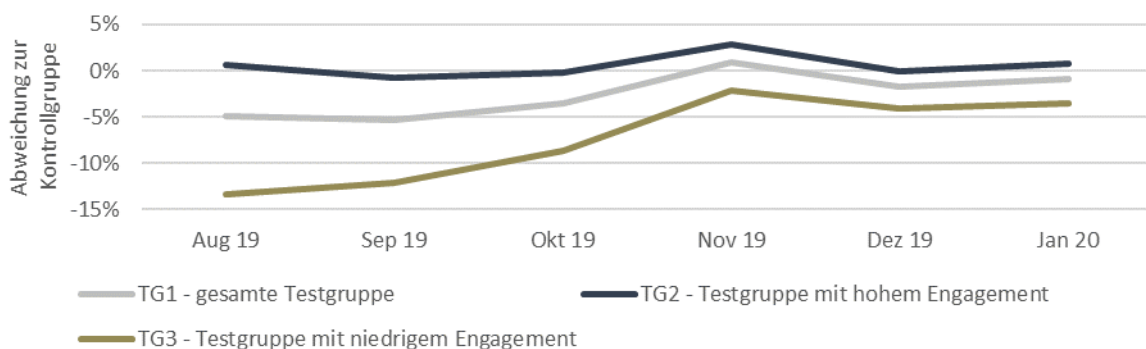


Abbildung 6: Prozentuale Abweichung des Gesamtverbrauchs der Testgruppe von der Kontrollgruppe

Während die Kunden mit niedrigem Engagement über den betrachteten Zeitraum unter dem Verbrauch der Kontrollgruppe blieb, lag der Verbrauch der Testkunden mit hohem Engagement im November 2,15% über dem Verbrauch der Kontrollgruppe. Tabelle 2 gibt einen Gesamtüberblick über die Mittelwerte (M) und Standardabweichungen (SD) der Verbrauchsdaten für alle Gruppen mit ihrer Stichprobengröße (n) im Zeitintervall August 2019 bis Januar 2020. Diese dienen als Grundlage für die

weitere statistische Analyse durch t-Tests. Eine statistische Betrachtung des erweiterten Zeitintervalls von Januar 2019 bis Januar 2020 ist in Anhang D angefügt.

Tabelle 2: Statistischer Überblick über die Gesamtverbrauchsdaten im Zeitraum August 2019 bis Januar 2020

Gruppe	n	Aug 19		Sep 19		Okt 19		Nov 19		Dez 19		Jan 19	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
KG	1037	7,75	4,31	8,14	4,45	8,70	4,81	9,43	5,39	9,92	6,02	9,70	5,72
TG1	88	7,36	3,92	7,71	4,19	8,38	4,87	9,51	6,69	9,76	6,25	9,61	6,36
TG2	53	7,79	4,31	8,08	4,47	8,67	5,23	9,70	6,78	9,92	6,10	9,78	6,27
TG3	35	6,71	3,17	7,16	3,72	7,95	4,32	9,23	6,62	9,52	6,55	9,35	6,57

Abkürzungen: KG = Kontrollgruppe, TG1 = Testgruppe gesamt, TG2 = Testgruppe mit hohem Engagement, TG3 = Testgruppe mit niedrigem Engagement, n = Stichprobenanzahl, M = arithmetisches Mittel, SD = Standardabweichung

Ergebnisse der t-Tests

Um den Einfluss der Reports auf den Gesamtstromverbrauch statistisch zu untersuchen, werden t-Tests zwischen der Kontrollgruppe und allen Ausprägungen der Testgruppe durchgeführt. Zuvor wird die Voraussetzung der Varianzhomogenität für die t-Tests untersucht. Eine Auswertung der Messdaten von Kontrollgruppe und Testgruppen für jeden Monat mit Hilfe von Levene-Tests ergab, dass Varianzhomogenität in allen Fällen gegeben ist (Signifikanz > 0,05, vgl. dazu Anhang D).

Zusammenfassend werden die t-Tests für unabhängige Stichproben einseitig auf einem Konfidenzniveau von 0,95 durchgeführt. Tabelle 3 zeigt die Ergebnisse der t-Tests für den Zeitraum August 2019 bis Januar 2020. Beispielhaft wird der Monat November 2019 nachfolgend detaillierter ausgewertet. Anschließend folgt eine zusammenfassende Beurteilung der Ergebnisse der t-Tests.

Tabelle 3: Ergebnisse der t-Tests für die Monate August 2019 bis Januar 2020

Monat	TG1			TG2			TG3		
	T	df	p	T	df	p	T	df	p
Aug. 19	-0,8109	1123	0,2088	0,07333	1088	0,5292	-1,4096	1070	0,0795
Sep. 19	-0,8775	1123	0,1902	-0,1031	1088	0,4589	-1,2974	1070	0,0974
Okt. 19	-0,5844	1123	0,2795	-0,0331	1088	0,4868	-0,2728	24,718	0,7873
Nov. 19	0,13114	1123	0,5522	0,34408	1088	0,6346	-0,2137	1070	0,4154
Dez. 19	-0,2445	1123	0,4035	-0,0099	1088	0,4961	-0,3848	1070	0,3502
Jan. 20	-0,1431	1123	0,4431	0,0934	1088	0,5372	-0,3492	1070	0,3635

Abkürzungen: TG1 = Testgruppe gesamt, TG2 = Testgruppe mit hohem Engagement, TG3 = Testgruppe mit niedrigem Engagement, T = T-Wert, df = Anzahl der Freiheitsgrade, p = Signifikanz

Im November 2019 ist der Verbrauch der Testgruppe (M = 9,51; SD = 6,69) im Vergleich zur Kontrollgruppe (M = 9,43; SD = 5,39) höher ausgefallen. Allerdings war die Differenz statistisch nicht signifikant ($t(1123) = 0,13$; $p > 0,05$). Der Verbrauch der Testgruppe mit hohem Engagement (M = 9,63; SD = 7,38) lag ebenfalls über dem Verbrauch der Kontrollgruppe. Auch diese Differenz war nicht signifikant ($t(1088) = 0,34$; $p > 0,05$). Daneben war der Gesamtverbrauch der Testgruppe mit niedrigem Engagement (M = 9,18; SD = 4,44) geringer als der Verbrauch der Kontrollgruppe. Auch diese Differenz war geringfügig ($t(1070) = -0,21$; $p > 0,05$).

Werden die Ergebnisse der t-Tests über den gesamten Zeitraum betrachtet, kann keine Differenz zwischen den Gruppen als signifikant eingestuft werden ($p > 0,05$). In den Monaten August und September 2019 war die Differenz jedoch signifikanter als während der Laufzeit des Experiments. Es kann nicht

festgestellt werden, dass der monatliche Report den Stromverbrauch der Testkunden gesenkt hat. Die in Kapitel 3 (S. 7) aufgestellte Forschungshypothese H1 kann somit nicht bestätigt werden.

3.2.2 Qualitative Ergebnisse

Neben der Untersuchung der Gesamtverbrauchsdaten konnten einige qualitative Erkenntnisse gewonnen werden. So betrug die Öffnungsrate der E-Mail mit dem Report als Anhang durchgehend über 75%. Darüber hinaus konnten die Testkunden auf freiwilliger Basis weitere Informationen zu ihrem Haushalt (z.B. Wohnungsgröße und Anzahl der Personen) teilen, um ihre Vergleichsberechnung entsprechend anpassen zu lassen. Dies taten insgesamt 62 Testkunden. Für diese Kunden wurde also während des Experimentes die Berechnung der normativen Vergleiche im Report auf diese angepasst.

Als Abschluss des Experiments konnten die Testkunden zu der gesamten Laufzeit in einer Online-Umfrage Feedback geben. Die Umfrage wurde am 7. Februar 2020 an die 88 Teilnehmer der Testgruppe geschickt. Insgesamt nahmen 20 Personen an der Umfrage teil.

Bei der Frage „Was hat dir am besten gefallen?“ wurden drei Themencluster identifiziert. Acht Personen (40%) gaben an, dass ihnen die historischen Vergleiche am besten gefallen haben. 35% antworteten, dass der Vergleich zu ähnlichen Haushalten, also der normative Vergleich, besonders gut war. Für fünf Teilnehmer (25%) war die allgemeine Übersicht, also die Zusammenstellung der verschiedenen Informationen im Report, hilfreich.

Darüber hinaus hat der Report laut Umfrageergebnissen 60% der Teilnehmenden geholfen, ihr Verbrauchsverhalten zu verändern. Als konkrete Änderungen wurden beispielsweise ein bewussterer Umgang mit dem Standby-Verbrauch, das Umstellen von Geräteeinstellungen auf energieeffizientere Modi oder auch kürzeres Duschen aufgeführt. Bei der Frage „Was willst du uns sonst noch in Bezug auf das Feature-Testing mitteilen?“ gaben sieben Personen an, dass sie sich über eine Umsetzung der Funktionalität in der App freuen würden. Allerdings wurde auch angemerkt, dass der Vergleich mit ähnlichen Haushalten transparenter erklärt werden könnte und dass der Report keine echte Neuerung zur Fresh Energy App sei.

4 Diskussion der Ergebnisse

Die Ergebnisse des Experiments wurden den Erwartungen an den Report nicht gerecht. Entsprechend der Forschungshypothese wurde erwartet, dass der Report den Stromverbrauch der Testgruppe gegenüber dem Verbrauch einer Kontrollgruppe senkt. Dies konnte statistisch nicht nachgewiesen werden. Der durchschnittliche Gesamtstromverbrauch der Testgruppe lag vor Beginn des Experiments (im September 2019) etwa 6% unter dem Verbrauch der Kontrollgruppe. Im November 2019 verbrauchte die Testgruppe dann sogar 0,85% mehr als die Kontrollgruppe – eine Steigerung von knapp 7% über zwei Monate. Danach sank der Verbrauch in den Monaten Dezember und Januar im Vergleich zur Kontrollgruppe wieder.

Auch die Aufteilung der Testgruppe durch die Variable „Engagement“ zeigt in der Auswertung des Gesamtstromverbrauchs keinen nennenswerten Effekt. Tatsächlich ist der durchschnittliche Verbrauch von TG2 durchgehend höher als der Verbrauch der TG3, was allerdings durch einen größeren Anteil mit verbrauchsstarken Haushalten zu erklären ist. In Relation zur Kontrollgruppe ist die Entwicklung der Testgruppen im Zeitverlauf des Experiments allerdings sehr ähnlich. Die Abweichungen der Verbrauchsstruktur sind indes vor und während des Experimentes bei einem Konfidenzniveau von 0,95 statistisch nicht signifikant.

Andererseits deuten einige Aspekte darauf hin, dass die Testkunden den Report als sinnvoll erachtet haben. Dafür spricht zum einen die konstant hohe Öffnungsrate der E-Mails von etwa 80%. Außerdem

gaben in der abschließenden Online-Umfrage 60% der Teilnehmer an, dass der Report das Stromverbrauchsverhalten positiv beeinflusst habe. Aus subjektiver Sicht der Kunden hatte der Report demnach einen Nutzen.

Somit kann keine verlässliche Aussage darüber getroffen werden, inwiefern der Verlauf der Verbrauchsdaten eine Aussagekraft besitzt oder ob der Report sogar für den höheren Verbrauch der Testgruppe im November verantwortlich ist. Auch kann kein signifikanter Unterschied zwischen den Testkunden, die den Report regelmäßig verwendet haben, und den Testkunden, die den Report nicht regelmäßig verwendet haben, festgestellt werden.

Entgegen der Erwartungen der Forschung konnte also nicht gezeigt werden, dass der Report den Stromverbrauch der Testgruppe senkt. Dafür könnte es einige Ursachen geben. Zum einen könnte das BEE-Potential der Testgruppe bereits durch die Fresh Energy App erschlossen sein. Eventuell ist der zusätzliche Anreiz in Form des Reportes nicht genug, um weitere BEE-Potentiale zu erschließen. Daneben spielt die überschaubare Datenbasis eine Rolle, insbesondere im Hinblick auf die Heterogenität der Haushaltsgruppen. So betrug etwa die Anzahl der Haushalte im Opower Experiment mehr als 20.000 [2, 22].

Zum anderen hat die Wahl der Auswertungsmethodik einen Einfluss auf die Ergebnisse. Es wurde vereinfachend angenommen, dass externe und saisonale Effekte sich in gleichem Maße auf die Kontroll- und Testgruppe auswirken. Es könnte daher zu Verzerrungen in den Ergebnissen kommen, falls die Test- und Kontrollgruppe unterschiedlich stark von diesen Effekten beeinflusst werden. Mit einer anderen Auswertungsmethodik könnten diese Verzerrungen vermieden werden.

Darüber hinaus könnten der inhaltliche Aufbau und die Gestaltung des Reports ausschlaggebend für die Ergebnisse sein. Eine neue Konfigurierung des Reports durch inhaltliche Schwerpunkte, eine andere Anordnung der einzelnen Elemente oder eine andere semantische Gestaltung könnte zu einem anderen Ergebnis führen. Fraglich ist insgesamt, welche konkrete Zusammenstellung aller zu berücksichtigten Feedbackansätze den optimalen Einfluss auf das Verhalten der Haushalte hat. Insbesondere das Zusammenspiel unterschiedlicher Feedback-Komponenten könnte eine zusätzliche unbekannte Größe darstellen.

5 Fazit und Ausblick

Die vorliegende Arbeit hat sich ausführlich mit der Anwendung von Smart Meter Daten zur Reduktion des Stromverbrauchs beschäftigt.

In einem Experiment wurde der Einfluss eines monatlichen Reports auf den Stromverbrauch einer Testgruppe im Vergleich zu einer Kontrollgruppe untersucht. Dabei konnte keine signifikante Änderung im Verbrauchsverhalten im Hinblick auf den Gesamtstromverbrauch belegt werden. Auch die deskriptive Analyse der disaggregierten Verbrauchsdaten zeigte keinen eindeutigen Einfluss des Reports. Allerdings wurde festgestellt, dass etwa 80% aller Testkunden den Report regelmäßig verwendeten. 73% der Testkunden waren sogar auf freiwilliger Basis dazu bereit, weitere persönliche Informationen über ihre Lebensbedingungen zur Verfügung zu stellen, um die normativen Vergleichswerte zu verbessern. Darüber hinaus gaben 60% der Teilnehmer der abschließenden Online-Umfrage an, dass der Report ihnen dabei geholfen hat, das Stromverbrauchsverhalten zu ändern. Es ist jedoch anzunehmen, dass die Umfrageergebnisse mit Verzerrungen verbunden sind. Auf Grundlage der quantitativen Auswertung der Stromverbrauchsdaten konnte dieser Effekt des Reports jedoch nicht bestätigt werden. Ein realisiertes Einsparpotenzial konnte auf Basis der erstellten Methodik demnach nicht nachgewiesen werden.

Insofern ist dieses Working Paper ein Plädoyer für die Anwendung der Verhaltensökonomik in der Energiewirtschaft. Sie zeigt auf, wie gemeinsam mit den Endkonsumenten geeignete Produkte auf Grundlage von Smart Meter Daten erstellt werden können. Das Working Paper zeigt gleichsam, wie sich die individuelle Wahrnehmung von Verhaltensanpassungen nicht notwendigerweise in quantitativ-messbare Effekte niederschlagen muss. Insgesamt konnte somit die Relevanz von BEE quantitativ nicht unterstrichen werden. Es wird deutlich, wie wichtig das experimentelle Design bei der Untersuchung von BEE Potentialen ist. Insbesondere die Bedeutung einer ausreichend großen Stichprobenanzahl und einer belastbaren statistischen Modellierung ist für weiterführende Forschungen hervorzuheben.

Zudem reiht sich die Arbeit in die bisher durchgeführten Studien zu Feedback basierend auf Smart Meter Daten ein. In Deutschland nimmt die Untersuchung sogar einen relativ einzigartigen Platz ein, da bundesweit bisher keine Untersuchungen zu Smart Metern Daten mit sekundlicher Auflösung und disaggregierten Daten erfolgt sind.

Limitation der wissenschaftlichen Arbeit

Die wissenschaftliche Arbeit ist in vielerlei Hinsicht limitiert. Die relativ kleine Stichprobengröße und die dadurch begrenzte Menge an Messdaten sind ein Faktor, der die Auswertung der Ergebnisse erschwert hat. Es kann zudem davon ausgegangen werden, dass die Nutzer von Fresh Energy zu den sogenannten „Early Adoptern“ zählen und somit von Grund auf sensibler für das Thema Energieeffizienz sind. Dies könnte die Theorie begründen, dass das BEE-Potential bei den Kunden von Fresh Energy bereits abgeschöpft ist. Andererseits könnten diese Personen auch empfänglicher für den Report gewesen sein, was sich in dem unerwartet hohen Engagement der Kunden widerspiegelt. Außerdem war die Dauer des Experiments mit insgesamt vier Monaten relativ kurz, sodass lediglich kurzfristige Effekte betrachtet werden konnten.

Darüber hinaus wird das Einsparpotential lediglich auf einer aggregierten Ebene betrachtet. Eine individuelle Modellierung des Einsparverhaltens einzelner Haushalte wird nicht behandelt. Ebenso ist die Untersuchung der disaggregierten Verbrauchsdaten auf eine deskriptive Analyse beschränkt. Die Verknüpfung der Ergebnisse mit den Kontrollvariablen der Gruppen durch eine explorative Clusteranalyse hat den Rahmen der vorliegenden Arbeit überschritten, könnte aber potentiell einen zusätzlichen Mehrwert bieten.

Außerdem besteht insgesamt eine Diskrepanz zwischen dem technologischen Standard von Smart Metern, welche durch den Smart Meter Rollout in Deutschland ausgebracht werden, und der Technologie von Fresh Energy. Dies stellt eine zukünftige Anwendbarkeit der erstellten Methodik für BEE-Potentiale in Frage, da in dieser Arbeit mit hochfrequenten Messdaten und Disaggregationsalgorithmen gearbeitet wurde. Zum jetzigen Zeitpunkt ist nicht abzusehen, wie sich die Anwendung der Messtechnik in Deutschland weiterentwickelt. Somit könnten die technischen Voraussetzungen für den breitflächigen Einsatz von BEE-Maßnahmen basierend auf Smart Meter Daten schlichtweg nicht gegeben sein.

Weiterführende Forschung

Die vorliegende Arbeit liefert einige Ansatzpunkte für zukünftige Forschungsleistungen. Zum einen wurden in der Auswertung monatliche Verbrauchsdaten verwendet, sodass lediglich ein aggregierter Einspareffekt untersucht wurde. In einem nächsten Schritt könnten kurzfristige bzw. wochentagspezifische Effekte untersucht werden, indem Messdaten kurz nach Erhalt von Feedback analysiert werden. Dazu wäre die Verwendung stündlicher oder täglicher Messdaten geeignet. Außerdem ist eine Untersuchung sinnvoll, welche die Entwicklung des Verbrauchsverhaltens der beobachteten Teilnehmer über einen längeren Zeitraum ermittelt. Hier ist fraglich, ob das Interesse der Verbraucher dauerhaft gehalten werden kann und welche Anreize langfristig zu energieeffizienterem Verhalten führen.

Zum anderen kann der Effekt einzelner verhaltensökonomischer Anreize weiter untersucht werden. Beispielsweise wäre eine Untersuchung der Integration von normativen Vergleichen (beispielsweise in der Fresh Energy App) denkbar. Durch die Änderung des Kommunikationsmediums von einem E-Mail Versand zur Integration in einer mobilen Anwendung ist eine veränderte Adaption bei den Kunden wahrscheinlich. In Bezug auf Fresh Energy könnte jede Neuerung der App zu einem anderen Konsumverhalten der Kunden führen. Somit wäre prinzipiell ein experimentelles Design für die zukünftige Produktentwicklung denkbar, bei der der Einfluss jeder Änderung auf das Verbrauchsverhalten untersucht werden könnte. Allerdings müssten hierfür auch überlagernde Effekte von anderen externen Einflussfaktoren berücksichtigt werden.

Im Kontext des deutschen Smart Meter Rollouts erscheinen insbesondere Studien notwendig, welche die Auswirkung von mMe und iMSys auf das Verbrauchsverhalten untersucht. Da der Smart Meter Rollout gerade mit zusätzlicher Transparenz und Verbrauchseinsparung wirbt, ist eine Überprüfung des realen Nutzens für private Haushalte sinnvoll. Wie in der Literaturrecherche gezeigt, ist insbesondere die Datenfrequenz ein vielversprechender Hebel, um BEE-Potentiale zu erschließen. Dabei ist es zu empfehlen, die privaten Haushalte proaktiv miteinzubeziehen und Marktlösungen zu schaffen, die die Kundenbedürfnisse erfüllen.

6 Danksagung

Diese Untersuchung wurde im Rahmen des Einsparzählerprogramms des BAFA (Förderkennzeichen: ESZ-A-2017031) erstellt, und gliedert sich in die Forschungsarbeiten des Verbundprojekts „Demand-Regio“ welches durch das Bundesministerium für Wirtschaft und Energie (BMWi) gefördert wird (Förderkennzeichen: 03ET4040C). Die Autoren möchten für die Finanzierung des Projekts danken. Des Weiteren gilt allen Kunden und Kundinnen von Fresh Energy ein Dankeschön, die sich die Zeit für Feedback im Rahmen des Versuchs genommen haben sowie eingewilligt haben, ihre Daten bereitzustellen.

7 Literatur

- [1] J. Schleich *et al.*, „Smart metering in Germany and Austria – results of providing feedback information in a field trial“, Fraunhofer ISI, Working Paper Sustainability and Innovation S6, 2011.
- [2] H. Allcott, „Social Norms and Energy Conservation“, *Journal of Public Economics*, Jg. 95, Nr. 5, S. 1982–2095, 2011.
- [3] J. Kelly und W. Knottenbelt, „Does disaggregated electricity feedback reduce domestic electricity consumption? A systematic review of the literature“, 3. Mai 2016. [Online]. Verfügbar unter: <http://arxiv.org/pdf/1605.00962v2>.
- [4] K. Carrie Armel, A. Gupta, G. Shrimali und A. Albert, „Is disaggregation the holy grail of energy efficiency? The case of electricity“, *Energy Policy*, Jg. 52, S. 213–234, 2013, doi: 10.1016/j.enpol.2012.08.062.
- [5] R. H. Thaler, „Toward a positive theory of consumer choice“, *Journal of Economic Behavior & Organization*, Jg. 1, Nr. 1, S. 39–60, 1980. [Online]. Verfügbar unter: [https://doi.org/10.1016/0167-2681\(80\)90051-7](https://doi.org/10.1016/0167-2681(80)90051-7)
- [6] D. Kahnemann und A. Tversky, „Prospect Theory: An Analysis of Decision under Risk“, *Econometrica*, Jg. 47, Nr. 2, S. 263–292, 1979.
- [7] R. H. Thaler und H. M. Shefrin, „An Economic Theory of Self-Control“, *Journal of Political Economy*, Jg. 89, Nr. 2, S. 392–406, 1981.
- [8] T. O'Donoghue und M. Rabin, „Doing It Now or Later“, *American Economic Review*, Jg. 89, Nr. 1, S. 103–124, 1999.
- [9] I. Fisher, *The Theory of Interest: As Determined by Impatience To Spend Income and Opportunity To Invest It*. New York, New York, USA: Macmillan, 1930.

- [10] G. Charness und M. Rabin, „Understanding Social Preferences with Simple Tests“, *The Quarterly Journal of Economics*, Jg. 117, Nr. 3, S. 817–869, 2002.
- [11] M. J.J. Handgraaf, M. A. van Lidth de Jeude und K. C. Appelt, „Public praise vs. private pay: Effects of rewards on energy conservation in the workplace“, *Ecological Economics*, Jg. 86, S. 86–92, 2013, doi: 10.1016/j.ecolecon.2012.11.008.
- [12] B. S. Frey und R. Jegen, „Motivation Crowding Theory: A Survey of Empirical Evidence“, *Zurich IEEER Working Paper No. 26; CESifo Working Paper Series No. 245*, 2000. [Online]. Verfügbar unter: <https://ssrn.com/abstract=203330>
- [13] E. L. Deci, „Effects of externally mediated rewards on intrinsic motivation“, *Journal of Personality and Social Psychology*, Jg. 18, Nr. 1, S. 105–115, 1971, doi: 10.1037/h0030644.
- [14] P. M. Todd und G. Gigerenzer, „Bounding rationality to the world“, *Journal of Economic Psychology*, Jg. 24, Nr. 2, S. 143–165, 2003, doi: 10.1016/S0167-4870(02)00200-3.
- [15] A. Tversky und D. Kahnemann, „Judgement under Uncertainty: Heuristics and Biases“, *Science*, Jg. 185, Nr. 4157, S. 1124–1131, 1974.
- [16] S. S. Iyengar und M. R. Lepper, „When choice is demotivating: Can one desire too much of a good thing?“, *Journal of Personality and Social Psychology*, Jg. 79, Nr. 6, S. 995–1006, 2000, doi: 10.1037//0022-3514.79.6.995.
- [17] W. Samuelson und R. Zeckhauser, „Status Quo Bias in Decision Making“, *Journal of Risk and Uncertainty*, Jg. 1, Nr. 1, S. 7–59, 1988. [Online]. Verfügbar unter: <https://doi.org/10.1007/BF00055564>
- [18] C. Alós-Ferrer, S. Hügelschäfer und J. Li, „Inertia and Decision Making“ (eng), *Frontiers in psychology*, Jg. 7, Nr. 169, 2016, doi: 10.3389/fpsyg.2016.00169.
- [19] I. P. Levin, S. L. Schneider und G. J. Gaeth, „All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects“, *Organizational Behavior and Human Decision Processes*, Jg. 76, Nr. 2, S. 149–188, 1998.
- [20] A. Tversky und D. Kahnemann, „The Framing of Decisions and the Psychology of Choice“, *Science*, Jg. 211, Nr. 4481, S. 453–458, 1981.
- [21] C. Fischer, „Feedback on household electricity consumption: a tool for saving energy?“, *Energy Efficiency*, Nr. 1, S. 79–104, 2008, doi: 10.1007/s12053-008-9009-7.
- [22] H. Allcott und T. Rogers, „The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation“, *American Economic Review*, Jg. 104, Nr. 10, S. 3003–3037, 2014, doi: 10.1257/aer.104.10.3003.
- [23] G. W. Hart, „Nonintrusive appliance load monitoring“, *Proc. IEEE*, Jg. 80, Nr. 12, S. 1870–1891, 1992, doi: 10.1109/5.192069.
- [24] T. Zia, D. Bruckner und A. Zaidi, „A hidden Markov model based procedure for identifying household electric loads“ in *IECON 2011 - 37th Annual Conference of IEEE Industrial Electronics*, Melbourne, Vic, Australia, 2011, S. 3218–3223, doi: 10.1109/IECON.2011.6119826.
- [25] M. J. Johnson und A. S. Willsky, „Bayesian Nonparametric Hidden Semi-Markov Models“, *Journal of Machine Learning Research*, Jg. 14, S. 673–701, 2013.
- [26] D. Srinivasan, W. S. Ng und A. C. Liew, „Neural-Network-Based Signature Recognition for Harmonic Source Identification“, *IEEE Trans. Power Delivery*, Jg. 21, Nr. 1, S. 398–405, 2006, doi: 10.1109/TPWRD.2005.852370.
- [27] A. Zoha, A. Gluhak, M. A. Imran und S. Rajasegarar, „Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey“ (eng), *Sensors (Basel, Switzerland)*, Jg. 12, Nr. 12, S. 16838–16866, 2012.
- [28] Umweltbundesamt, „Entwicklung von Methoden zur Evaluierung von Energieeinsparung: Ermittlung des Energieeffizienzfortschritts in Deutschland gemäß der EU-Richtlinie 2006/32/EG“, 2011. [Online]. Verfügbar unter: <http://www.uba.de/uba-info-medien/4076.html>. Zugriff am: 1. März 2020.

A. Anhang: Statistischer Überblick der Kontrollvariablen

Statistische Aufbereitung der numerischen Variable „Alter“:

	Kontrollgruppe		TG1		TG2		TG3	
	M	SD	M	SD	M	SD	M	SD
Alter [y]	44,43	12,45	42,36	12,56	40,51	11,08	43,58	13,41

Abkürzungen: TG1 = Testgruppe gesamt, TG2 = Testgruppe mit hohem Engagement, TG3 = Testgruppe mit niedrigem Engagement, M = arithmetisches Mittel, SD = Standardabweichung

Statistische Aufbereitung der nominalen Kontrollvariablen:

Variable	Kontrollgruppe (n = 1037)		TG1 (n = 88)		TG2 (n = 35)		TG3 (n = 53)	
	Anzahl	Anteil	Anzahl	Anteil	Anzahl	Anteil	Anzahl	Anteil
Geschlecht = männlich	837	80,71%	74	84,09%	27	77,14%	47	88,68%
Geschlecht = weiblich	200	19,29%	14	15,91%	8	22,86%	6	17,14%
PLZ = 0	67	6,46%	6	6,82%	2	5,71%	4	11,43%
PLZ = 1	195	18,80%	12	13,64%	8	22,86%	4	11,43%
PLZ = 2	106	10,22%	9	10,23%	4	11,43%	5	14,29%
PLZ = 3	116	11,19%	10	11,36%	5	14,29%	5	14,29%
PLZ = 4	139	13,40%	17	19,32%	13	37,14%	4	11,43%
PLZ = 5	111	10,70%	12	13,64%	9	25,71%	3	8,57%
PLZ = 6	88	8,49%	4	4,55%	3	8,57%	1	2,86%
PLZ = 7	72	6,94%	6	6,82%	1	2,86%	5	14,29%
PLZ = 8	84	8,10%	4	4,55%	1	2,86%	3	8,57%
PLZ = 9	59	5,69%	8	9,09%	7	20,00%	1	2,86%

Abkürzungen: TG1 = Testgruppe gesamt, TG2 = Testgruppe mit hohem Engagement, TG3 = Testgruppe mit niedrigem Engagement, n = Stichprobenanzahl

B. Anhang: Monatlicher Report

Fresh Review: Dein September 2019




Dein Verbrauch lag bei
355,54 kWh und 112,93 €.



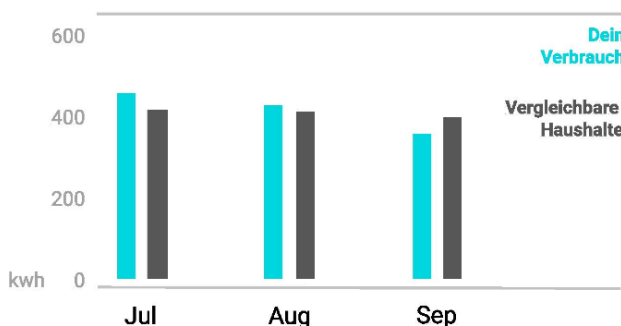
68,11 kWh und 19,01 €
weniger als im letzten Monat.

Das hat sich bei deinen Geräten verändert


	August 2019		September 2019	Vergleichbare Haushalte
☾ Stand By	95,95 kWh	↗	129,92 kWh	101,78 kWh
🧊 Kühlschrank	12,72 kWh	↗	43,4 kWh	36,98 kWh
💡 Beleuchtung	21,93 kWh	→	21,36 kWh	40,09 kWh
🧺 Waschmaschine	16,84 kWh	↘	11,51 kWh	19,14 kWh
🧼 Spülmaschine	22,83 kWh	↘	6,14 kWh	16,23 kWh
🔥 Ofen	2,67 kWh	↘	0,3 kWh	4,12 kWh
🔌 Andere Geräte	225,07 kWh	↘	142,91 kWh	181,33 kWh


Dein CO² Fußabdruck entspricht etwa*  457 km oder  801 km oder  2646 km

Dein Verbrauch im Vergleich mit anderen Haushalten



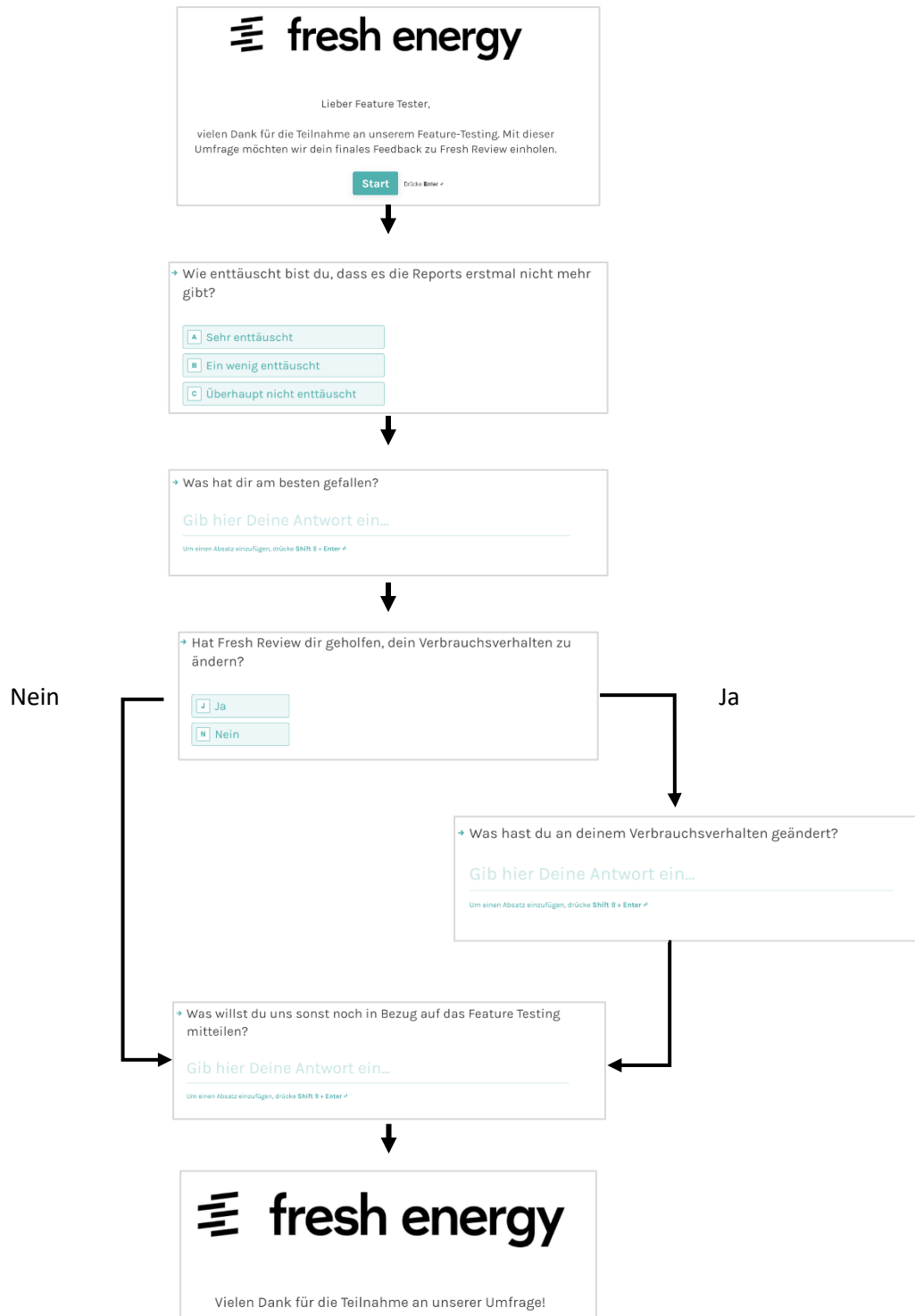
💡 Wusstest du schon, dass...

 ...ein Großteil des Energiebedarfs einer Waschmaschine bei der Erhitzung des Wassers entsteht? Wer seine Wäsche mit 30°C wäscht, spart bis zu 70% Strom gegenüber einer Wäsche mit 60°C.

 ...jedes Jahr ca. 4 Mrd. € Kosten für den Stand-By Verbrauch in Deutschland entstehen? Wer seine Geräte richtig ausschaltet, kann durchschnittlich mehr als 100€ pro Jahr sparen.

*Jede Kilowattstunde erzeugt in Deutschland aktuell durchschnittlich 474 g CO² - auch wenn du Ökostrom bezieht.

C. Anhang: Fließschema der abschließenden Online-Umfrage



D. Anhang: Statistische Aufbereitung der Gesamtverbrauchsdaten

Gesamtverbrauch, detaillierte Statistik der monatlichen, durchschnittlichen Tagesverbräuche im Zeitraum Januar 2019 bis Januar 2020 für die final gebildeten Gruppen.

Monat	Kontrollgruppe			TG1			TG2			TG3		
	M*	SD*	n	M*	SD*	n	M*	SD*	n	M*	SD*	n
Jan. 19	10,12	6,07	881	8,72	4,51	72	9,33	4,69	45	7,71	4,08	27
Feb. 19	9,63	5,56	910	8,49	4,36	79	8,76	4,43	47	8,11	4,3	32
Mrz. 19	9,27	4,95	944	8,61	4,88	80	9,18	5,36	48	7,76	3,98	32
Apr. 19	8,61	4,61	978	7,86	4,31	80	8,35	4,56	48	7,13	3,87	32
Mai 19	8,48	4,64	1000	7,86	4,23	83	8,34	4,49	49	7,16	3,77	34
Jun. 19	7,99	4,43	1031	7,57	4,38	87	8,27	4,86	52	6,54	3,35	35
Jul. 19	7,81	4,35	1033	7,21	4,12	88	7,88	4,71	53	6,19	2,76	35
Aug. 19	7,75	4,31	1037	7,36	3,92	88	7,79	4,31	53	6,71	3,17	35
Sep. 19	8,14	4,45	1037	7,71	4,19	88	8,08	4,47	53	7,16	3,72	35
Okt. 19	8,7	4,81	1037	8,38	4,87	88	8,67	5,23	53	7,95	4,32	35
Nov. 19	9,43	5,39	1037	9,51	6,69	88	9,7	6,78	53	9,23	6,62	35
Dez. 19	9,92	6,08	1037	9,76	6,25	88	9,92	6,1	53	9,52	6,55	35
Jan. 20	9,7	5,72	1037	9,61	6,36	88	9,78	6,27	53	9,35	6,57	35

Abkürzungen: TG1 = Testgruppe gesamt, TG2 = Testgruppe mit hohem Engagement, TG3 = Testgruppe mit niedrigem Engagement, n = Stichprobenanzahl, M = arithmetisches Mittel, SD = Standardabweichung

* Werte in [kWh / d], gerundet

Ergebnisse der Levene-Tests auf Varianzhomogenität zur Kontrollgruppe:

Monat	TG1		TG2		TG3	
	F	Signifikanz*	F	Signifikanz*	F	Signifikanz*
Aug. 19	0,9805	0,3223	0,0484	0,826	1,7864	0,1817
Sep. 19	0,5553	0,4563	0,1082	0,7423	0,8207	0,3652
Okt. 19	0,1327	0,7157	0,029	0,8649	0,1517	0,697
Nov. 19	1,0762	0,2998	0,8317	0,362	0,3098	0,5779
Dez. 19	0,0051	0,9432	0,003	0,9565	0,0076	0,9307
Jan. 19	0,0043	0,9475	0,0103	0,9193	0,0001	0,9927

Abkürzungen: TG1 = Testgruppe gesamt, TG2 = Testgruppe mit hohem Engagement, TG3 = Testgruppe mit niedrigem Engagement;

*Eine Signifikanz > 0,05 bedeutet, dass Varianzhomogenität gegeben ist