

Kosten und Kostentreiber von unerwünschten Werbemails (Spam) – Eine empirische Analyse auf Provider- und Anwenderseite



Michel Clement, Dominik Papies, Harder-Johann Boie*

Überblick

- Etwa 71% des globalen E-Mail-Aufkommens sind mittlerweile unerwünschte Werbemails (Spam). Die Kosten, die durch diese Art des Direkt-Marketings entstehen, betreffen Organisationen zentral auf der IT-Ebene und dezentral auf der Mitarbeiterebene.
- Die Kosten auf IT-Ebene sind im Vergleich zu den Kosten auf Mitarbeiterebene, die vorrangig durch Arbeitszeitverluste entstehen, zu vernachlässigen.
- Unsere empirische Analyse unter 1.000 Universitätsmitarbeitern zeigt, dass die Kosten pro Mitarbeiter im Mittel bei 531 Euro (2004) bzw. 447 Euro (2005) pro Jahr liegen.
- Zentrale Kostentreiber auf Mitarbeiterebene sind neben der Anzahl der empfangenen Spam-Mails die Verbreitung der eigenen E-Mail-Adresse, die wahrgenommene Reaktanz auf die Spam-Belästigung und der dadurch entstehende Vertrauensverlust in das Kommunikationsmittel sowie die Zeit für die Überprüfung der E-Mail auf Spam.
- Während der Einsatz von Spam-Filtern kurzfristig keine Kostenersparnisse liefert, lassen sich in der langen Sicht signifikante Einsparungen erzielen.

Keywords Email-marketing · direkt marketing · spam · spam-filter · mixture regression

Eingegangen: 25. August 2005

JEL: M15, M31

Prof. Dr. Michel Clement (✉)

Lehrstuhl für Marketing und Medienmanagement, Institut für Marketing und Medien, Universität Hamburg, Von-Melle-Park 5, 20146 Hamburg. Telefon: 040 42838 4275. Email: michel.clement@uni-hamburg.de, URL: www1.uni-hamburg.de/imm

Dipl.-Kfm. Dominik Papies (✉)

Lehrstuhl für Marketing und Medienmanagement, Institut für Marketing und Medien, Universität Hamburg, Von-Melle-Park 5, 20146 Hamburg. Telefon: 040 42838 4157. Email: dominik.papies@uni-hamburg.de, URL: www1.uni-hamburg.de/imm

Dipl.-Kfm. Harder-Johann Boie (✉)

Lehrstuhl für Innovation, Neue Medien und Marketing, Institut für Innovationsforschung, Christian-Albrechts-Universität zu Kiel, Westring 425, 24098 Kiel. Telefon: 0431 880 1541. Email: boie@bwl.uni-kiel.de, URL: <http://www.bwl.uni-kiel.de/bwlinstitute/Innovation-Marketing/new/de/home>

A. Kostenrelevanz von Spam

Das Internet bietet eine Vielzahl von Möglichkeiten zur direkten Ansprache von (potenziellen) Kunden. So bekommen E-Mail-Nutzer täglich Angebote zur Verbesserung der Gesundheit bzw. der finanziellen Situation oder für Gewinne aus Glücksspielen. Sofern diese Angebote unerwünscht sind, wird diese Art des Direkt-Marketings allgemein als Spam¹ angesehen. Es lassen sich folgende Merkmale von Spam identifizieren (OECD, 2005): (1) Spam wird massenhaft und auf elektronischem Wege zugestellt (z. B. E-Mail oder Voice-over-IP), (2) Spam wird ohne Zustimmung des Empfängers versendet und ist unerwünscht, (3) Spam hat einen kommerziellen und werbenden Charakter, (4) die Zustellung von Spam kann kaum gestoppt werden, obwohl oft entsprechende Vorkehrungen zum Abbestellen vorgetäuscht werden und (5) die Absender von Spam sind häufig unbekannt oder nutzen gefälschte Adressen. Ob jedoch eine Mail im konkreten Fall als Spam angesehen wird, hängt von der subjektiven Wahrnehmung des Nutzers ab. So ist eine E-Mail von z. B. Dell für einige Personen informativ, während sie für andere Spam ist. Daher kann die Spam-Belastung nur direkt bei den Nutzern über eine individuelle Messung der Spam-Wahrnehmung quantifiziert werden.

Spam existiert, weil es ein profitables Geschäft ist (Sipior, Ward und Bonner, 2004). So kann bei Spam bereits eine sehr geringe Kaufquote von 0,0001 Prozent profitabel sein, da durch die digitale Distribution nur geringe Einrichtungs- und Versandkosten entstehen und die Preise für Millionen von Mailadressen immer geringer werden (Vircom, 2004). Zudem verfolgen Spammer mehrere Geschäftsmodelle. Zum einen können (möglichst validierte) E-Mail-Adressen an andere Spammer weiterverkauft werden. Zum anderen können Spammer direkt eigene Produkte bewerben. In den meisten Fällen werden fremde Produkte gegen Zahlung einer Provision offeriert. Die Gewinne variieren in Abhängigkeit vom Produkt, von der Zeitspanne der Spam-Kampagne und natürlich von der Öffnungs- bzw. Kaufquote, die wiederum von der inhaltlichen Qualität der E-Mails beeinflusst wird. Entsprechend gaben Spammer an, dass sie Gewinne zwischen 100 und 6.500 US \$ pro Woche erzielen (Vircom, 2004). Auch in Deutschland lohnt sich das Geschäft (BSA, 2004): So klicken 31% der Nutzer regelmäßig auf Links innerhalb der Werbe-Mails und 10% haben bereits auf diese Weise Waren erworben. Weitere 18% klicken zur Abbestellung der Spam-Mails auf den dafür vorgesehenen Link – mit dem Effekt, dass der Spammer seiner Datenbank eine weitere aktive E-Mail-Adresse hinzufügen kann. Mit zunehmender Erfahrung der Nutzer im Umgang mit Spam sinken zwar diese Quoten, gleichzeitig sinken aber auch die Preise je Adresse, so dass sich Spamming weiterhin lohnt. Seit Oktober 2003 machen E-Mails, die eindeutig als Spam identifiziert werden können, mehr als die Hälfte des globalen E-Mail-Verkehrs aus (Abbildung 1).

Entsprechend beklagen E-Mail-Provider die Belastung durch Spam (Topf, 2005). So meldet T-Online, dass täglich ca. eine Milliarde Spam-Mails empfangen werden (Kuri, 2006). Auch Organisationen bzw. Firmen werden zunehmend belastet. Das Rechenzentrum der Christian-Albrechts-Universität zu Kiel, an der diese Studie durchgeführt wurde, berichtet von einer Spam-Quote von 75% und mehr (Abbildung 2).²

Der Empfang von Spam hat sich längst von einer Belästigung zu einem volkswirtschaftlichen Problem entwickelt (OECD, 2005). Solange es international noch keine implementierten Richtlinien gegen Spam gibt, stehen Organisationen aus Produktivitäts-

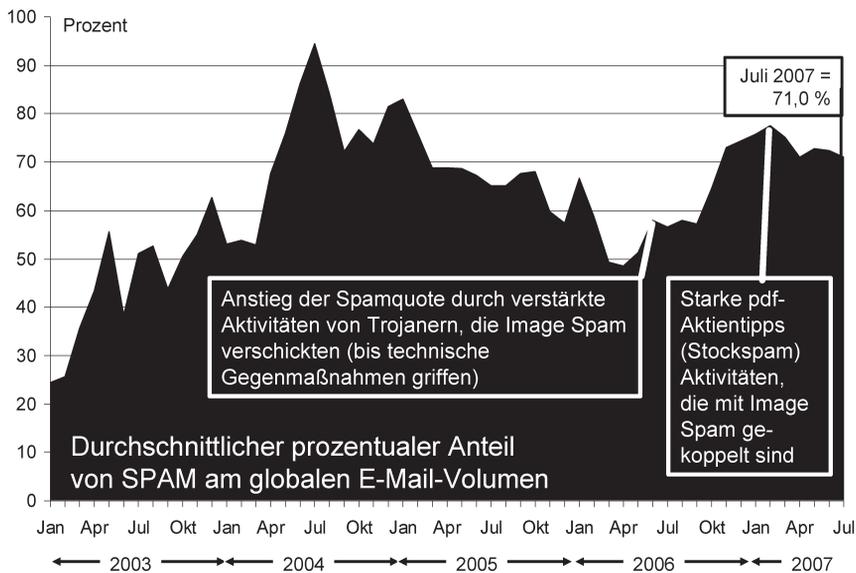


Abb. 1. Der Anteil von Spam am globalen E-Mailaufkommen (Quelle: Messagelabs.com)

	<u>E-Mail pro Monat</u>	<u>SPAM-Quote</u>	
Oktober 2003	1,8 Mio.	75 %	
Januar 2005	4,4 Mio.	86 %	+244 %
März 2006	4,3 Mio.*	88 %*	Quelle: Rechenzentrum (2006)

* Geänderte Berechnungsgrundlage. De fakto sind bereits vor der Annahme von Mails aufgrund von Ziel-Adress-Informationsprüfprozeduren (forward-callback) bzw. aufgrund der Absender-Adress-Informationsprüfung bereits ca. 1,4 Mio. E-Mails vor der eigentlichen Zustellung als SPAM identifiziert und abgewiesen worden. Demnach wären es bei der alten Berechnung ca. 5,7 Mio. E-Mails mit einer SPAM-Quote von 92% gewesen.

Abb. 2. Anteil von empfangenen Spam-Mails an der Christian-Albrechts-Universität zu Kiel

gründen unter dem Druck, mit eigenen Mitteln die durch Spam entstehenden Kosten einzudämmen. Daher ist das Interesse groß, zum einen zu bestimmen, welche Kosten genau entstehen und zum anderen zu analysieren, welche Einflüsse auf die Höhe der Kosten wirken, um geeignete Maßnahmen gegen Spam abzuleiten. Diesen beiden Fragestellungen widmet sich der vorliegende Aufsatz.

Bei einer Literaturdurchsicht wird deutlich, dass ein erheblicher Forschungsbedarf bei der Bestimmung der Kosten von Spam besteht (Joseph und Thevaranjan, 2005; Schryen, 2004). Management-Studien, die sich mit den Kosten von Spam beschäftigen, stellen sich meist als wenig fundiert heraus (z. B. Nucleus Research, o. V., 2004). Häufig wird von wenigen, allgemeinen Kennzahlen sowohl auf die Kosten der gesamten Organisation, als auch auf die des einzelnen Mitarbeiters geschlossen. Dabei werden die individuellen Kosten der Mitarbeiter nicht berücksichtigt. Auch werden die Kosten, die zentral anfallen zumeist ignoriert, so dass die Kosten von Spam nicht umfassend in die Analyse eingehen.

Aus diesem Grund ist ein Ansatz notwendig, der sowohl die Gesamtkosten von Spam auf Provider- als auch auf Anwenderseite erfasst. Dieser Aufsatz setzt sich daher zwei Ziele:

- Zum einen werden die Kosten von Spam in einer Organisation am Beispiel einer Universität bestimmt, wobei die Kosten von Spam auf Anwender- sowie auf Providerseite berücksichtigt werden. Hierfür wurden Expertengespräche mit den Verantwortlichen des Rechenzentrums geführt (zentrale Messung) und ein Online-Panel mit zwei Wellen unter den Mitarbeitern der Universität vorgenommen (dezentrale Messung).
- Zum anderen sollen die zentralen Kostenkomponenten identifiziert werden. Die Identifikation dieser Faktoren erlaubt eine zielgenauere Implementierung von Maßnahmen gegen Spam. Anhand eines Panels werden die zentralen Kostentreiber empirisch mit Hilfe von Mixture-Regression-Modellen untersucht.

Im Weiteren geben wir einen kurzen Überblick zum Stand der Forschung, um dann das eigene Untersuchungsdesign im Abschnitt C vorzustellen. Abschnitt D widmet sich der Analyse der Kosten auf Providerseite, während Abschnitt E die Kosten auf Anwenderseite betrachtet. Der Aufsatz schließt mit Schlussfolgerungen für Wissenschaft und Management.

B. Stand der Forschung

In Abbildung 3 wird der bisherige Stand der Forschung zu den Kosten von Spam dargestellt. Potenzielle und bestehende Kunden mit Werbung per E-Mail anzusprechen, ist eine effiziente Form des **Direkt-Marketings** (Dwyer, 1997; Krafft, 2002). Die direkte Ansprache des Kunden mittels E-Mail stellt hierbei eines der zentralen Instrumente im Marketing-Mix dar – insbesondere, wenn es sich um E-Commerce-Unternehmen handelt (Krafft und Peters, 2005). Die Untersuchungen zur Effizienz dieser Marketing-Maßnahmen betrachten allerdings nur selten den schmalen Grat zwischen direkter Kundenansprache, die vom Kunden gewünscht ist, und der unerwünschten Werbezustellung, die als Spam empfunden wird (Chu, Gerstner und Hess, 1995). So existieren kaum Untersuchungen dazu, inwieweit Direkt-Marketing-Maßnahmen bei Kunden zu Reaktanz führen (Akcura und Srinivasan 2005; McKnight, Choudhury und Kacmar, 2002). Vielmehr wird

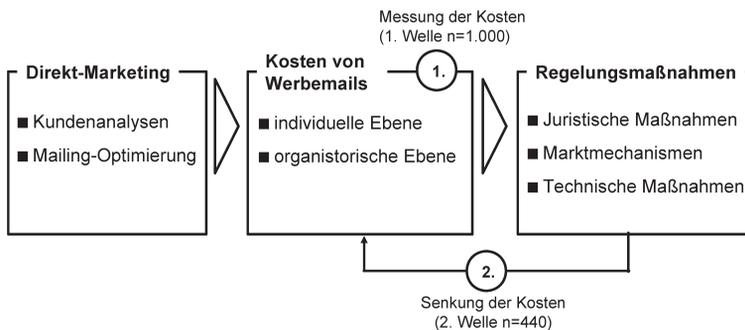


Abb. 3. Forschungsbereiche zu den Kosten von Spam

deutlich, dass dieser Forschungsbereich zumeist die direkte Kunden-Ansprache als vorteilhaft darstellt (Hesse, Krafft und Peters, 2005) – ein Grund für die zunehmende Spam-Belastung.

Einen Überblick über die wenigen bislang durchgeführten Studien zu den **Kosten von Spam** geben die Kommission der Europäischen Gemeinschaften (2004) und die OECD (2005). Dort wird zwischen direkten und indirekten Kosten unterschieden. Direkte Kosten beinhalten: (1) Arbeitszeitverluste beim Löschen von Spam-Mails durch E-Mail-Nutzer, (2) Kosten der IT-Verantwortlichen für die Recherche und Implementierung geeigneter Maßnahmen gegen Spam und (3) Kosten für den E-Mail-Provider, der wegen Spam mehr Bandbreite und Speicherkapazität zur Verfügung stellen muss (Schryen, 2006). Indirekte Kosten werden durch Fehlklassifikationen hervorgerufen. Weitere indirekte Kosten entstehen, wenn Spam Viren oder Würmer beinhaltet (Schryen, 2004). Dennoch wird eine substantielle Forschungslücke offenbar, denn bislang ist unklar, in welcher Höhe einer Organisation tatsächlich relevante Kosten entstehen und welche Faktoren die einzelnen Kostenarten maßgeblich beeinflussen.

Entsprechend interessieren sich immer mehr Forscher für **Regelungsmaßnahmen** zum Eindämmen der Spam-Flut. Es stehen drei Forschungsbereiche im Fokus, mit denen der Spam-Belastung entgegen getreten werden kann. Zum einen lassen sich *juristische Maßnahmen* anführen, die sich kurzfristig auszahlen (Sipior, Ward und Bonner, 2004). So sank die Spam-Belastung stets, wenn die rechtlichen Konsequenzen international verschärft wurden. Allerdings wirken nationale juristischen Maßnahmen zumeist nur so lange, bis die Spammer auf internationales Terrain ausgewichen sind, so dass keine langfristige Kostenrelevanz vorliegt (Zhang, 2005; OECD, 2005). Auch *ökonomische Ansätze* werden zunehmend diskutiert (Joseph und Thevaranjan, 2005; Kraut et al., 2005). Hierbei geht es vor allem darum, mittels Preismechanismen (Porto) oder Steuern die Kosten des unerwünschten Mailversands zu steigern. Unter Wohlfahrtsgesichtspunkten kann gezeigt werden, dass das so genannte „Bonded Sender Programm“ sehr geeignet ist (Joseph und Thevaranjan, 2005). Hierbei überwacht eine zentrale Instanz die Erlaubnis, dass Personen per E-Mail kontaktiert werden dürfen (Whitelist). Sofern eine Person eine andere Person kontaktieren möchte ohne auf dessen Whitelist zu stehen, muss diese Person bei der Instanz einen Geldbetrag hinterlegen. Wenn der Adressat nach Erhalt der E-Mail der Instanz meldet, dass der Inhalt Spam gewesen ist, dann behält die Instanz den zuvor eingezahlten Geldbetrag ein. Bislang haben sich diese (Markt-)Lösungen jedoch noch nicht durchgesetzt und es ist aufgrund des erheblichen internationalen Abstimmungsaufwands auch nicht zu vermuten, dass sich dies kurzfristig ändern wird (Zhang, 2005). Dem gegenüber stehen die *technischen Maßnahmen* (z. B. Bayes-Filter), die zentral im Rechenzentrum oder dezentral beim Nutzer eingesetzt werden und denen meist eine kostenreduzierende Wirkung zugeschrieben wird.³ Schryen (2004) stellt in seinem Überblicksaufsatz die Forschungsströme in der (Wirtschafts-) Informatik dar. Diese fokussieren sich zum einen auf Maßnahmen zum Blockieren von IP-Nummern, Filtermechanismen oder die Authentifizierung. Zum anderen werden in der Industrie und den Standardisierungsgremien Ansätze diskutiert, die mit einer Veränderung der prinzipiellen E-Mail-Kommunikationsstruktur einhergehen (z. B. Beschränkungen der Anzahl versendbarer E-Mails pro E-Mail-Konto und Tag). Offen bleibt hingegen, ob Kosten durch zentrale und dezentrale Spam-Filtermaßnahmen eingespart werden können oder ob die Kosten, die durch den Filter entstehen, überwiegen.

Zusammenfassend lässt sich konstatieren, dass die Forschung sich zunehmend der Spam-Problematik widmet. Erklärt wird vor allem, warum es sich lohnt, per E-Mail Werbung zu betreiben. Dementsprechend wird auch immer mehr auf die dadurch entstehenden Kosten abgestellt. Interessanterweise werden diese Kosten jedoch nur sehr grob quantifiziert und es bleibt offen, welche Einflussfaktoren die Kosten vorrangig treiben. Dieser Forschungslücke widmen wir uns in einer eigenen empirischen Untersuchung, in der auch die Forschungsfrage adressiert wird, inwieweit ein Spam-Filter die Kosten reduzieren kann.

C. Untersuchungsdesign

Als Untersuchungsobjekt dient die Christian-Albrechts-Universität zu Kiel. Ihre Größe entspricht mit ca. 8.000 Mitarbeitern der eines mittelgroßen Unternehmens. Zusätzlich stehen alle Bereiche der Organisation, in denen Kosten durch Spam anfallen können, zur Datenerhebung zur Verfügung: einerseits die Anwender, andererseits der E-Mail-Provider und die IT-Abteilung. Es können daher die gesamten Kosten von Spam in einer Organisation ermittelt werden. An der Universität fallen die beiden Bereiche E-Mail-Provider und IT-Abteilung im Rechenzentrum zusammen. Die Kosten von Spam müssen nur in zwei Bereichen erhoben werden. Dabei wird das Rechenzentrum im Folgenden als Provider, die Summe der Mail-Empfänger der Organisation als Endanwender bezeichnet. Kosten auf Providerseite entstehen z.B. durch das gestiegene Datentransfer-Volumen oder den Speicherbedarf auf dem Mail-Server. Werden auf Provider-Seite Maßnahmen gegen Spam eingeführt, treten für die Einrichtung und Wartung weitere Kosten auf. Bei den Endanwendern fallen Kosten an, die durch den Verlust von Arbeitszeit aufgrund der Bearbeitung von Spam entstehen. Außerdem müssen Finanzmittel aufgewendet werden, um eine Software gegen Spam zu erwerben. Bei den Mitarbeitern entstehen nicht nur quantifizierbare Kosten, sondern ebenfalls ein nicht monetärer Nutzenentgang, da die

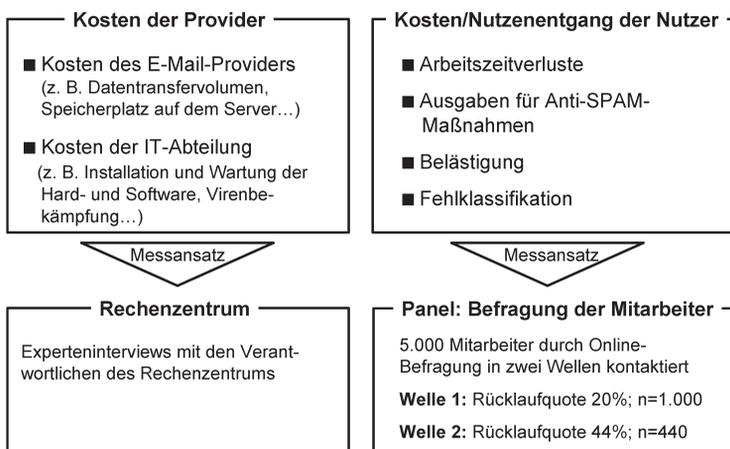


Abb. 4. Untersuchungsdesign

Mitarbeiter durch die Spam-Belästigung in ihrem Arbeitsfluss gestört werden. Zusätzlich müssen Endanwender – sofern sie Filter-Maßnahmen gegen Spam getroffen haben – Zeit für Kontrollen aufwenden, um so mögliche Fehlklassifikationen aufzudecken.

Die Kosten, die in beiden Bereichen der Organisation durch Spam-Mails entstehen, wurden im Rahmen einer empirischen Studie ermittelt. Es wurden Expertengespräche mit den verantwortlichen Fachkräften auf Seiten des Rechenzentrums geführt. Zudem wurden Universitätsmitarbeiter mit der Bitte angeschrieben, an einem Panel teilzunehmen und einen Online-Fragebogen auszufüllen (Abbildung 4). Mit diesen Daten können die Kosten auf Organisationsebene zu mehreren Zeitpunkten bestimmt werden. Die erste Welle der Befragung fand im Oktober 2004 und die zweite Welle der Befragung im November 2005 statt. Die Befragten der ersten Welle hatten unmittelbar nach der Befragung die Möglichkeit, ihren Spam-Filter zu verschärfen, so dass in der zweiten Welle gemessen wurde, ob die Maßnahmen angepasst wurden und wie stark die Kosten reduziert werden konnten.

Die Bestimmung der Kosten in Form der Arbeitszeitverluste stellt die Basis (d.h. die abhängige Variable) dar, um darauf aufbauend die zentralen Kostentreiber (per Regressionsanalyse) zu ermitteln. Die alternativen Kostenmaße (Ausgaben für Anti-Spam-Maßnahmen, Belästigung und Fehlklassifikation) werden zur Validierung herangezogen.

Die Systematisierung der Einflussfaktoren der Kosten von Spam auf Mitarbeiterebene basiert auf den Erkenntnissen der Theorien, die auf der „Theory of Reasoned Action“ aufbauen. (d.h. Theorie des geplanten Verhaltens und Technology Acceptance Model; Ajzen und Fishbein, 1977; Armitage und Conner, 2001 sowie Lwin und Williams, 2003; Abbildung 5). Die dahinter liegende Idee ist zum einen, dass die Adoption von neuen Technologien weniger durch produktspezifische sondern vielmehr durch individuen-spezifische Faktoren erklärt werden kann. Zum anderen wird angenommen, dass das zukünftige

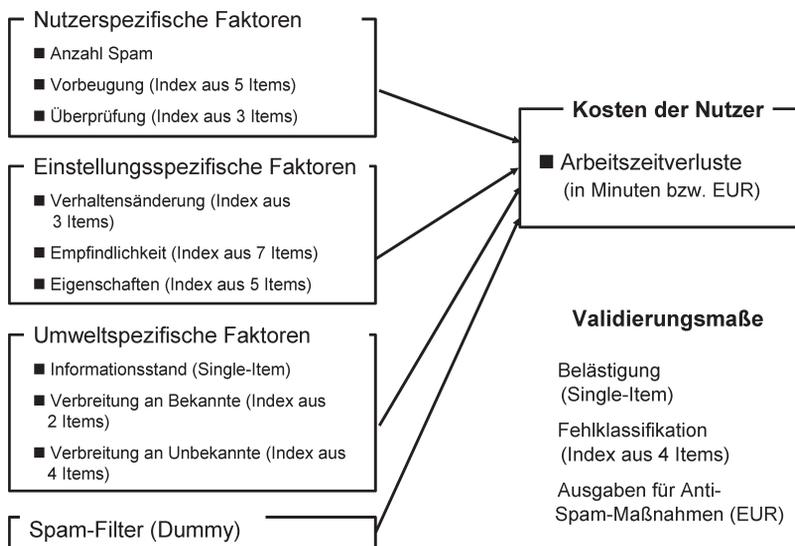


Abb. 5. Modellierung der Kosten auf Mitarbeiterseite

tige Verhalten eine Funktion der Verhaltensabsicht ist, welche wiederum u.a. durch die Kosten beeinflusst wird. Das bedeutet, je höher die durch Spam-Mails entstehenden Kosten sind, desto wahrscheinlicher wird ein Mitarbeiter eine Maßnahme gegen Spam ergreifen, da der wahrgenommene Nutzen der Gegenmaßnahme steigt (Davis, Bagozzi und Warshaw, 1989). Da im Kontext von Spamfiltern der Einfluss des sozialen Drucks als gering eingestuft werden kann, wurde auf die explizite Messung verzichtet. Die einzelnen Faktoren setzen sich aus verschiedenen Indizes zusammen. Die zugrunde liegende Operationalisierung der Indizes wurde so weit wie möglich in Anlehnung an bisherige Konstruktoperationalisierungen vorgenommen. Wir folgen bei der Messung der Konstrukte Albers und Hildebrandt (2006), die zu einer formativen Messung raten, sofern nach der Wirkung der Faktoren geforscht wird, die durch das Management beeinflusst werden können. Da reflektive Items ein Konstrukt widerspiegeln, es jedoch (in Sinne von Inputfaktoren) nicht formen, lassen sich keine konkreten Management-Empfehlungen ableiten. Da wir jedoch an Managementimplikationen interessiert sind, formulieren wir unsere Skalen mit formativen Indikatoren (Diamantopoulos und Winklhofer, 2001; Jarvis, Mackenzie und Podsakoff, 2003). Dies impliziert zugleich, dass die Items, die das Konstrukt formen (z. B. Spam-Empfindlichkeit) nicht notwendigerweise miteinander korreliert sein müssen (so wie bei reflektiven Messungen, die zumeist anhand korrelationsbasierter Maße evaluiert werden; Rossiter, 2002). Weder formative noch reflektive Indikatoren sind bislang für derartige Fragestellungen dokumentiert worden, so dass eine Validierung der Skalen nur auf der Basis der Inhaltsvalidität erfolgen kann (Rossiter, 2005).⁴ Die Stärke der Einflüsse der einzelnen Indizes auf die Kostenkomponenten wird mit Regressionsanalysen untersucht. Nachdem die Vorgehensweise dargestellt wurde, werden im Folgenden die Ergebnisse auf Seiten des Providers (Abschnitt D) bzw. auf Seiten der Anwender (Abschnitt E) erörtert.

D. Analyse der Kosten auf Providerseite

Die in diesem Abschnitt verwendeten Daten entstammen mündlichen und schriftlichen Befragungen der IT-Experten des universitätsinternen Rechenzentrums aus dem Februar 2005.⁵ Seit April 2003 stellt das Rechenzentrum seinen Nutzern eine zentrale Maßnahme gegen Spam zur Verfügung. Alle aus dem Internet eingehenden E-Mails werden auf Spam-verdächtige Merkmale untersucht, entsprechend markiert und an die Mailserver weitergeleitet. Im Februar 2005 erreichten täglich ca. 170.000 E-Mails das Rechenzentrum, die auf Spam (und Viren) untersucht und an die Empfänger weitergeleitet wurden. Hinzu kamen ca. 20.000 E-Mails, die innerhalb des Universitätsnetzes versendet und nicht auf Spam überprüft wurden.

Bevor an der Universität mit der technischen Umsetzung einer Maßnahme gegen Spam begonnen werden konnte, mussten verschiedene Gremien über die geplanten Maßnahmen informiert und deren Zustimmung eingeholt werden.⁶ Die Anschaffungskosten für einen Anti-Spam-Server beziffert das Rechenzentrum auf 3.000 Euro. Der Erwerb der Software für den Anti-Spam-Server (SpamAssassin) verursachte hingegen keine Kosten, da es sich um Open-Source-Software handelt. Die einmaligen Kosten beim Aufbau der zentralen Maßnahme gegen Spam sind in Tabelle 1 aufgeführt. Die Einführung und

Tab. 1. Kosten auf Providerseite (Stand: Oktober 2004)

Einmalige Tätigkeiten	Zeitbedarf	Kosten
Administrative Tätigkeiten (z. B. Abstimmung von Maßnahmen mit Datenschutzbeauftragten, Vorstellung möglicher Maßnahmen bei Vorgesetzten etc).	58 Stunden	1.740 €
Anschaffungskosten Hard- / Software	–	3.000 €
Aufbau der Infrastruktur (z. B. Einarbeitungszeit, Beschaffung, Installation etc.)	78 Stunden	2.340 €
Laufzeitabhängige Kosten auf Provider-Seite	Zeitbedarf	Kosten
Wartung und Weiterentwicklung der Gegenmaßnahmen, Support(anfragen) etc.	268 Stunden	8.040 €
SUMME Gesamtkosten auf Providerseite	404 Stunden	15.120 €

Betreuung der Anti-Spam-Maßnahmen oblag hauptsächlich Personen der Gehaltsstufe BAT II A. Anhand dieser Information konnten die aufgewendeten Arbeitszeiten in Euro überführt werden.⁷

Die Gesamtkosten auf Providerseite belaufen sich demnach auf rund 15.000 € und sind damit eher zu vernachlässigen (Tabelle 1). Dieses Ergebnis verdeutlicht, dass die Kosten, die den Providern entstehen, nicht unbedingt in den Fokus gerückt werden sollten.

E. Analyse der Kosten auf Anwenderseite

I. Stichprobe

Die Kosten durch Spam auf Seiten der Endnutzer wurden über zwei Online-Befragungen der Mitarbeiter an der Christian-Albrechts-Universität zu Kiel ermittelt, die jeweils im Oktober 2004 und 2005 durchgeführt wurden. Zu diesen Zeitpunkten nutzten etwa 8.000 Mitarbeiter die E-Mail-Infrastruktur des Rechenzentrums. Etwa 5.000 dieser Nutzer konnten im Oktober 2004 direkt per E-Mail kontaktiert werden. Der andere Teil konnte aufgrund der zersplitterten Netzinfrastruktur nicht erreicht werden, da die Subnetze durch eigene Administratoren verwaltet werden.⁸ Von diesen 5.000 Personen beantworteten exakt 1.000 Personen (20%) den Online-Fragebogen vollständig. Im Oktober 2005 wurden die 1.000 Teilnehmer der ersten Befragung per E-Mail zu einer zweiten Umfrage eingeladen, an der 440 Personen (44%) teilnahmen.⁹ In dieser Befragung wurden die Kostenmaße erneut erhoben, um so Veränderungen über die Zeit zu messen. Zudem wurde erfragt, ob die Teilnehmer ihre Maßnahmen gegen Spam seit der letzten Befragung verschärft haben.

II. Operationalisierungen und deskriptive Ergebnisse

1. Wahrgenommener Spam-Anteil

In der ersten Welle wurden knapp 45% der E-Mails von den Mitarbeitern als Spam angesehen (Abbildung 6). Da die Universitäten bereits in den Anfangsjahren des Internet online waren und die Mail-Adressen auf den Uni-Webseiten typischerweise gelistet sind,

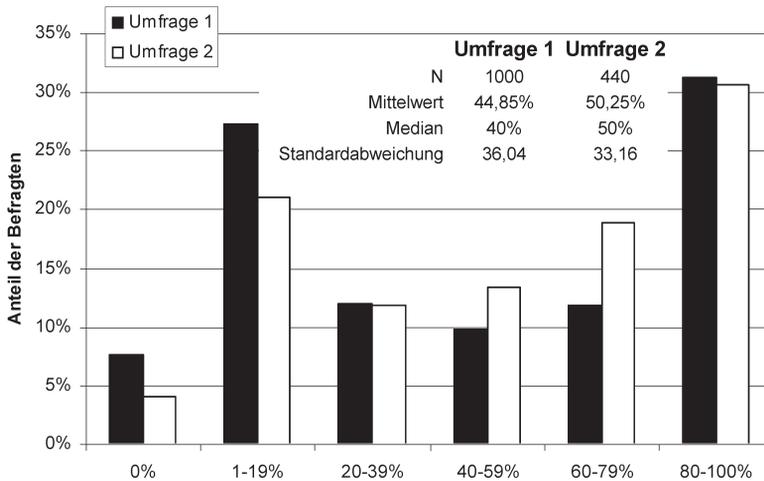


Abb. 6. Anteil von Spam am gesamten E-Mail-Volumen der Befragten

sind diese hohen Werte durchaus plausibel. Es zeigt sich zudem, dass der wahrgenommene Spam-Anteil in der zweiten Welle signifikant angestiegen ist (50%; vgl. auch Tabelle 5). Des Weiteren fällt auf, dass die Spam-Belastung heterogen ist und auf zwei oder mehr latente Segmente geschlossen werden kann. Abbildung 6 verdeutlicht, dass eine große Anzahl von Mitarbeitern wenig bzw. gar kein Spam empfängt. Ein Blick auf das Ende der Skala offenbart jedoch, dass es auch Mitarbeiter gibt, die mehr als 80% ihrer E-Mails als Spam betrachten. Die Ursache hierfür kann sowohl in der tatsächlich unterschiedlich verteilten Anzahl von Spam-Mails als auch in der heterogen verteilten subjektiven Definition der Spam-Eigenschaften liegen. Diese Verteilung ist für die nachfolgenden Analysen von hoher Relevanz, da Verfahren zu wählen sind, die diese Heterogenität kontrollieren können (z. B. Mixture-Regression-Modelle).

2. Operationalisierung und Messung der Kosten und des Nutzenentgangs durch Spam

Abbildung 5 zeigt mehrere Kostenarten im Zusammenhang mit Spam auf. Im Folgenden wird vorrangig auf die Kosten durch Arbeitszeitverluste eingegangen, da diese zum einen in Expertengesprächen als die wichtigste Komponente angesprochen wurden und zum anderen auch Gegenstand der Managementstudien waren (OECD, 2005). Die anderen Kostenmaße dienen im Weiteren der Validierung der Ergebnisse der Arbeitszeitverluste. Durch Spam-Mails entstehen bei den Mitarbeitern der Universität **Arbeitszeitverluste**, die an mehreren Stellen im Panel erfasst wurden. Zum einen müssen Spam-Mails aus den empfangenen E-Mails mindestens täglich¹⁰ manuell aussortiert oder die automatisch durch Gegenmaßnahmen erfolgte Sortierung auf Richtigkeit überprüft werden (Tabelle 2). Zum anderen gab ungefähr die Hälfte aller Befragten in der ersten Welle an, eine Maßnahme gegen Spam einzusetzen. Die Suche nach einer geeigneten Maßnahme sowie deren Installation und Konfiguration hat zusätzliche Arbeitszeit gekostet, zum Teil nicht nur die eigene, sondern auch die eines technisch versierten Mitarbeiters. Diese Zeitauf-

Tab. 2. Verlorene Zeit durch Spam auf Endnutzerseite

		STUDIE 1 (Total)	STUDIE 1 (Panel)	STUDIE 2 (Panel)
	Maß- einheit	N=1000 Mean (Stdv)	N=440 Mean (Stdv)	N=440 Mean (Stdv)
Arbeitszeitverluste pro Jahr	Minuten	1247,22	1214,25	1078,61²
Index = (w₁·250)+w₂+w₃+w₄+w₅+w₆+w₇		(1686,33)	(1527,53)	(1359,43)
(w ₁) Zeit für tägliche Spam-Bearbeitung	Minuten	4,87 (6,70)	4,73 (6,07)	4,20 ² (5,39)
(w ₂) Zeit für Recherche Rechenzentrum (RZ)-Filter	Minuten	17,03 (20,11)	18,36 (19,59)	16,56 ¹ (14,98)
(w ₃) Zeit für Einrichtung RZ-Filter	Minuten	15,68 (18,87)	16,41 (20,11)	19,09 ^{ns} (30,15)
(w ₄) Zeit für Einrichtung RZ-Filter durch einen Mitarbeiter	Minuten	11,73 (11,63)	11,32 (11,40)	9,69 ^{ns} (7,37)
(w ₅) Zeit für Recherche nach einer alternativen Gegenmaßnahme	Minuten	53,34 (54,83)	55,59 (48,87)	40,61 ^{ns} (39,69)
(w ₆) Zeit für Installation / Konfiguration einer alternativen Gegenmaßnahme	Minuten	34,40 (72,23)	32,19 (51,17)	22,4 ^{ns} (41,20)
(w ₇) Zeit für Installation / Konfiguration einer alternativen Gegenmaßnahme durch einen Mitarbeiter	Minuten	23,37 (21,24)	22,46 (20,70)	35,59 ¹ (74,07)

¹: Mittelwert ist auf dem Niveau von 10% von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

²: Mittelwert ist auf dem Niveau von 5% von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

^{ns}: Mittelwert ist nicht von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

wendungen wurden in sechs weiteren Variablen erfasst und beschreiben, so die Annahme, einen einmaligen Vorgang. Alle sieben Variablen werden in einem Index zu der abhängigen Variable „Arbeitszeitverlust“ zusammengefasst, die in Minuten gemessen wird. Im Mittel verlieren alle 1000 Teilnehmer der ersten Welle jedes Jahr 1.247 Minuten Arbeitszeit durch Spam. Betrachtet man nur die Werte der 440 Befragten, die an beiden Befragungen teilgenommen haben, so zeigt sich, dass die Kosten über die Zeit signifikant gesunken sind (1.214 Minuten versus 1.078 Minuten), jedoch die Höhe der Zeitverluste auch im Oktober 2005 noch substanziell ist.

Da die Besoldungsstufen und das Alter der Mitarbeiter erfragt wurden, können die Arbeitszeitverluste mit dem Brutto-Lohn (ohne Nebenkosten) gewichtet und in Euro überführt werden (Tabelle 3). Es zeigt sich, dass die Kosten der Panelteilnehmer von 531 Euro (Okt. 2004) auf 447 Euro (Okt. 2005) sanken. Dennoch – multipliziert man die Kosten von 447 Euro mit der Anzahl der Mitarbeiter der Universität (5.000) so entsteht ein Scha-

Tab. 3. Verlorenes Gehalt durch Spam auf Endnutzerseite

		STUDIE 1 (Total)	STUDIE 1 (Panel)	STUDIE 2 (Panel)
	Maß- einheit	N=1000 Mean (Stdv)	N=440 Mean (Stdv)	N=440 Mean (Stdv)
Arbeitszeitverluste pro Jahr	Euro	538,36	530,97	447,10¹
Index = (w₁·250+w₂+w₃+w₄+w₅+w₆+w₇) · Gehalt		(849,27)	(787,22)	(598,18)

¹: Mittelwert ist auf dem Niveau von 5% von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

den durch Spam in Höhe von 2,235 Mio. Euro pro Jahr.¹¹ Das Ergebnis verdeutlicht, dass die Kosten auf Seiten der Mitarbeiter deutlich höher sind als auf Providerseite.

Neben den Zeitverlusten lässt sich Spam auch durch alternative Kostenmaße erfassen: Spam kann schlicht stören und zu einer **Belästigung** werden. Diese „psychologischen Kosten“ haben wir direkt mittels einer 5-Punkt-Likert-Skala erfragt (Rossiter, 2002). Tabelle 4 zeigt, dass die Belästigung durch Spam im Oktober 2004 im Mittel bei 3,4 und somit im oberen Bereich der Skala lag. In 2005 konnte dieser Wert signifikant reduziert werden (auf 3,2) – er ist jedoch noch immer hoch. Der Nutzenentgang durch **Fehlklassifikationen** von E-Mails durch Spam-Filter wird mit vier Variablen erfasst, die zu einem Index summiert werden. Die Sorge, eine wichtige E-Mail (auch „Ham“ genannt) irrtümlich als Spam anzusehen und zu löschen, liegt im mittleren Bereich der Skala. Die Unsicherheit, eine wichtige E-Mail wegen möglicher Anti-Spam-Maßnahmen zu verlieren, ist deutlich höher. Es zeigt sich auch, dass diese Kosten über die Zeit von 10,5 auf 11,4 signifikant gestiegen sind.

Tab. 4. Alternative Kostenmaße

	Maßeinheit	STUDIE 1	STUDIE 1	STUDIE 2
		(Total)	(Panel)	(Panel)
		N=1000	N=440	N=440
		Mean (Stdv)	Mean (Stdv)	Mean (Stdv)
Belästigung	Likert (1-5)	3,444	3,414	3,225¹
		(1,442)	(1,451)	(1,389)
Fehlklassifikation	Likert (4-20)	10,359	10,570	11,470 ¹
Index = $w_1 + w_2 + w_3 + w_4$		(4,331)	(4,391)	(4,663)
(w_1) Sorge: Lösche irrtümlich erwünschte E-Mail	Likert (1-5)	2,644	2,714	2,998 ¹
		(1,372)	(1,370)	(1,410)
(w_2) Fremde E-Mail kam nicht an	Likert (1-5)	2,094	2,084	2,339 ¹
		(1,397)	(1,411)	(1,507)
(w_3) Sorge: Erhalte erwünschte E-Mail nicht	Likert (1-5)	3,000	3,084	3,220 ¹
		(1,386)	(1,388)	(1,390)
(w_4) Sorge: Eigene E-Mail kommt nicht an	Likert (1-5)	2,621	2,689	2,914 ¹
		(1,342)	(1,335)	(1,356)

¹: Mittelwert ist auf dem Niveau von 5% von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

Auch durch den **Kauf einer Anti-Spam-Software** können Kosten entstehen. Da aber gute Anti-Spam-Maßnahmen größtenteils kostenlos verfügbar sind, wurden insgesamt nur 536 Euro von elf Befragten für diesen Zweck ausgegeben. Aufgrund der geringen Relevanz wird im Weiteren darauf verzichtet, diese Variable zu betrachten.

Die Ergebnisse des Panels zeigen, dass die Kosten auf Mitarbeiterebene sehr hoch sind und im Laufe der Zeit sowohl die Arbeitszeitverluste wie auch die wahrgenommene Belästigung sinken während die Sorge vor Fehlklassifikationen ansteigt. Im Folgenden werden die potenziellen Kostentreiber dargestellt.

3. Operationalisierung und Messung der Kostentreiber

Zunächst werden die **nutzerspezifischen Faktoren** diskutiert (Tabelle 5). Der erste Block (Spam-Menge) beschreibt die E-Mail-Nutzung. Im Mittel bekommen die Mitarbeiter der

Tab. 5. Nutzerspezifische Faktoren

	Maßeinheit	STUDIE 1 (Total) N=1000 Mean (Stdv)	STUDIE 1 (Panel) N=440 Mean (Stdv)	STUDIE 2 (Panel) N=440 Mean (Stdv)
Spam-Menge Spam = (w₁·w₂) / 100	Anzahl	16,880 (28,397)	17,838 (29,574)	16,345^{ns} (24,514)
(w ₁) Anzahl der empfangenen E-Mails pro Tag	Anzahl	23,378 (31,615)	24,501 (32,631)	23,305 ^{ns} (27,592)
(w ₂) Spam-Anteil	Prozent	44,846 (36,038)	44,052 (35,972)	50,252 ¹ (33,155)
Spam-Vorbeugung Index = (w₁+w₂+w₃+w₄+w₅) / 5	Likert (1-5)	2,256 (0,916)	2,218 (0,903)	2,369¹ (0,928)
(w ₁) Veröffentlichung der E-Mail-Adresse auf Webseiten vermieden	Likert (1-5)	2,942 (1,692)	2,874 (1,679)	3,275 ¹ (1,635)
(w ₂) Weitergabe der E-Mail-Adresse vermieden	Likert (1-5)	2,539 (1,458)	2,550 (1,462)	2,614 ^{ns} (1,457)
(w ₃) Verwendung einer alternativen E-Mail-Adresse	Likert (1-5)	2,723 (1,833)	2,658 (1,830)	2,705 ^{ns} (1,822)
(w ₄) Verwendung einer ungewöhnlichen E-Mail-Adresse	Likert (1-5)	1,257 (0,796)	1,211 (0,688)	1,309 ¹ (0,848)
(w ₅) Entfernung von E-Mail-Liste(n) beantragt	Likert (1-5)	1,821 (1,437)	1,798 (1,421)	1,943 ¹ (1,525)
Spam-Überprüfung Index = (w₁ + w₂ + w₃) / 3	Likert (1-5)	3,708 (0,544)	3,684 (0,569)	-
(w ₁) Begutachtung des Absenders	Likert (1-5)	4,733 (0,768)	4,697 (0,848)	-
(w ₂) Begutachtung der Betreff-Zeile	Likert (1-5)	4,721 (0,783)	4,706 (0,832)	-
(w ₃) Öffne E-Mails	Likert (1-5)	1,670 (0,961)	1,650 (0,963)	-

¹: Mittelwert ist auf dem Niveau von 5% von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

^{ns}: Mittelwert ist nicht von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

Universität 23 E-Mails pro Tag, wovon knapp 45% (2004) bzw. 50% (2005) als Spam angesehen werden. Die Menge der Spam-Mails unterscheidet sich nicht signifikant in den beiden Wellen. Im zweiten Bereich (Spam-Vorbeugung) wird ermittelt, ob die Befragten hinsichtlich der Vorbeugung von Spam aktiv waren. Wir gewichten hierbei die Maßnahmen gleich und führen sie zu einem Index zusammen. Es wird deutlich, dass die Befragten im Laufe der Zeit signifikant mehr Maßnahmen durchgeführt haben, um Spam vorzubeugen. Der letzte Block (Spam-Überprüfung) zeigt auf, wie mutmaßliche Spam-Mails kontrolliert werden. Je größer der Wert dieser Index-Variablen ist, desto mehr Zeit wird für die Überprüfung der E-Mails aufgebracht. Diese Frage wurde nur in der ersten Welle gestellt, da wir annehmen, dass sich diese Variablen nicht über die Zeit verändern.¹²

Die **einstellungsspezifischen Faktoren** und ihre Kerngrößen werden in Tabelle 6 dargestellt. Die Belastung durch Spam hat die Befragten in beiden Wellen nicht dazu bewegt, die Nutzung von E-Mails zu reduzieren (Index: Verhaltensänderungen). Auch das generelle Vertrauen in E-Mails als Kommunikationsmittel ist durchaus hoch und

Tab. 6. Einstellungsspezifische Faktoren

	Maß- einheit	STUDIE 1 (Total)	STUDIE 1 (Panel)	STUDIE 2 (Panel)
		N=1000 Mean (Stdv)	N=440 Mean (Stdv)	N=440 Mean (Stdv)
Verhaltensänderungen				
Index = (w₁+w₂+w₃) / 3	Likert (1-5)	2,126 (0,885)	2,120 (0,888)	2,020² (0,791)
(w ₁) E-Mail-Nutzung wegen Spam reduziert	Likert (1-5)	1,276 (0,775)	1,251 (0,758)	1,230 ^{ns} (0,688)
(w ₂) E-Mail-Bearbeitungszeit wegen Spam erhöht	Likert (1-5)	2,807 (1,499)	2,872 (1,532)	2,707 ² (1,322)
(w ₃) Vertrauen in E-Mail wegen Spam gesunken	Likert (1-5)	2,294 (1,282)	2,235 (1,277)	2,125 ¹ (1,190)
Spam-Empfindlichkeit				
Index = (w₁+w₂+w₃+w₄+w₅+w₆+w₇) / 7	Likert (1-5)	3,256 (0,623)	3,232 (0,640)	-
(w ₁) Werbe-E-Mails von unbekanntem Absender	Likert (1-5)	4,809 (0,617)	4,789 (0,657)	-
(w ₂) Unbestellte E-Mails von politischer/juristischer Gruppierung	Likert (1-5)	4,568 (0,870)	4,561 (0,898)	-
(w ₃) Unbestellte E-Mails von nicht kommerzieller Organisation	Likert (1-5)	4,410 (1,010)	4,386 (1,050)	-
(w ₄) E-Mails mit Witzen, Geschichten, Links	Likert (1-5)	2,594 (1,455)	2,555 (1,442)	-
(w ₅) Werbe-E-Mails von Geschäftspartnern	Likert (1-5)	2,561 (1,346)	2,550 (1,364)	-
(w ₆) E-Mails mit großen Dateien (Bilder, Videos)	Likert (1-5)	2,141 (1,334)	2,084 (1,320)	-
(w ₇) E-Mails kommerzieller Art mit Erlaubnis	Likert (1-5)	1,711 (1,061)	1,699 (1,064)	-
Spam-Eigenschaften				
Index = (w₁+w₂+w₃+w₄+w₅) / 5	Likert (1-5)	4,030 (0,805)	4,045 (0,825)	-
(w ₁) Spam-Mails wurden nicht angefordert	Likert (1-5)	4,520 (0,925)	4,538 (0,926)	-
(w ₂) Spam-Flut ist nicht zu stoppen	Likert (1-5)	4,209 (1,122)	4,228 (1,123)	-
(w ₃) Schadenspotenzial der Spam-Mails für den eigenen Computer	Likert (1-5)	4,185 (1,254)	4,167 (1,267)	-
(w ₄) Wahrgenommene Menge der empfangenen Spam-Mails	Likert (1-5)	3,835 (1,353)	3,836 (1,361)	-
(w ₅) Zeitbedarf zur Bearbeitung der Spam-Mails	Likert (1-5)	3,399 (1,385)	3,455 (1,403)	-

¹: Mittelwert ist auf dem Niveau von 10% von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

²: Mittelwert ist auf dem Niveau von 5% von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

^{ns}: Mittelwert ist nicht von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

nimmt mit der Zeit zu. Es wird deutlich, dass die Verhaltensänderungen in der zweiten Welle signifikant geringer sind als im Oktober 2004. Als zweiter Index wird die Spam-Empfindlichkeit einbezogen. Die Nutzer unterscheiden sich in ihrer Empfindlichkeit bezüglich der Wahrnehmung von Spam. Daher bilden wir den ungewichteten Mittelwert hinsichtlich der Bewertung unterschiedlicher Mail-Inhalte (z. B. Witzmails) als Index für die Empfindlichkeit. Wir gehen davon aus, dass auch dieser Index sich nicht innerhalb eines Jahres verändert. Dies gilt ebenso für den letzten Index, der die wahrgenommenen störenden Eigenschaften von Spam misst.¹³ Hierbei wird der Mittelwert über die möglicherweise als störend empfundenen Eigenschaften von unerwünschten Werbemails gebildet.

Tabelle 7 zeigt die **umweltspezifischen Faktoren**. Während die Befragten ihren Spam-Wissensstand in der ersten Welle als durchschnittlich einschätzen, weist er in der zweiten Welle ein signifikant höheres Niveau auf. Dies kann auf eigene aber auch auf Lerneffekte, die durch diese Studie entstanden, zurückzuführen sein. Letztlich ist es aber unerheblich für die Kostenanalyse, woher die Informationen stammen – vielmehr ist zu analysieren, ob und wie stark ein gesteigertes Informationsniveau die Kosten von Spam senken kann. Spammer nutzen häufig E-Mail-Adressen, die öffentlich bekannt sind. Daher ist der Verbreitungsgrad einer E-Mail-Adresse ein möglicher Einflussfaktor, den es zu kontrollieren gilt. Wir unterscheiden die Verbreitung der E-Mail-Adresse an Bekannte und an Unbekannte, indem wir einen Index als Summe mehrerer nominalskaliertes Variablen bilden. Wir gehen davon aus, dass das Niveau der beiden letztgenannten Variablen über die Zeit konstant bleibt.

Tab. 7. Umweltspezifische Faktoren

	Maß- einheit	STUDIE 1 (Total)	STUDIE 1 (Panel)	STUDIE 2 (Panel)
		N=1000 Mean (Stdv)	N=440 Mean (Stdv)	N=440 Mean (Stdv)
Informationsstand zu Spam	Likert (1-5)	2,699 (1,121)	2,672 (1,106)	3,087¹ (0,999)
Verbreitung der E-Mail-Adresse an Bekannte Index = w_1+w_2	0-2	1,460 (0,606)	1,466 (0,603)	–
(w_1) an Arbeitskollegen/Unternehmensvertreter weitergegeben	0/1	0,916 (0,278)	0,918 (0,274)	–
(w_2) an Freunde / Bekannte weitergegeben	0/1	0,544 (0,498)	0,548 (0,498)	–
Verbreitung der E-Mail-Adresse an Unbekannte Index = $w_1+w_2+w_3+w_4$	0-4	1,780 (0,965)	1,902 (0,972)	–
(w_1) auf Webseiten hinterlegt	0/1	0,721 (0,449)	0,732 (0,444)	–
(w_2) in einem Online-Verzeichnis hinterlegt	0/1	0,673 (0,469)	0,736 (0,441)	–
(w_3) in Internet-Foren angegeben	0/1	0,069 (0,254)	0,093 (0,291)	–
(w_4) bei der Anmeldung auf Internetseiten / zu Newslettern angegeben	0/1	0,317 (0,466)	0,341 (0,475)	–

¹: Mittelwert ist auf dem Niveau von 5% von Studie 1 (Panel) verschieden (t-Test bei gepaarten Stichproben)

Bei den **technologiespezifischen Faktoren** fokussieren wir uns auf den Einsatz von Schutzmaßnahmen gegen Spam. Ein Spam-Filter senkt möglicherweise die Kosten, da er eine Vorabklassifikation vornimmt (Schryen, 2004). Allerdings können die Kosten auch steigen, wenn erwünschte Mails fälschlicherweise herausgefiltert werden. In der vorliegenden Stichprobe setzen in der ersten Welle 527 Personen einen oder mehrere Filter ein. Der Anteil im Panel ($n=440$) beträgt in der ersten Welle 51,6% und in der zweiten Welle 67,7%. Insofern ist die Nutzung von Spam-Filtern über die Zeit signifikant gestiegen. Dies kann als Indiz gelten, dass der Einsatz von Spam-Filtern ein erhebliches Kostensenkungspotenzial mit sich bringt. Zudem haben wir die Panelteilnehmer in der zweiten Welle gefragt, inwieweit sie die bereits getroffenen Maßnahmen innerhalb des letzten Jahres verstärkt haben. Insgesamt haben 42,3% der Panelteilnehmer den Spam-Schutz erhöht (Abbildung 7).

III. Regressionsanalysen zur Ermittlung der Kostentreiber auf Anwenderseite

Zur Analyse der Effektstärken werden Regressionsanalysen eingesetzt. Die abhängige Variable stellt die Arbeitszeitverluste in Minuten dar. Die unabhängigen Variablen weisen untereinander keine erheblichen Korrelationen auf und die jeweiligen VIF-Werte liegen in beiden Wellen unter 1,4.¹⁴ Auf das Einbeziehen einer Dummy-Variablen, die die Nutzung eines Spam-Filters berücksichtigt, wird verzichtet. Die Einbeziehung dieser Variablen in die Regressionen würde zu Endogenitätsproblemen führen, denn es ist unklar, ob die Nutzer aufgrund hoher Kosten einen Filter installieren oder ob der Filter hohe Kosten hervorruft.¹⁵ Wir diskutieren die Filterwirkung separat im Abschnitt IV.

Zunächst wird eine Regressionsanalyse (OLS) für alle 1000 Teilnehmer der ersten Welle und anschließend je eine Regressionsanalyse für die 440 Teilnehmer der beiden Wellen im Panel vorgenommen (Tabelle 8). Wir diskutieren vorrangig die Ergebnisse des Panels ($n=440$) dahingehend, welche Kostentreiber einen Erklärungsbeitrag für die Arbeitszeitverluste liefern und wie sich diese Werte über die Zeit verändert haben. Zur Kontrolle möglicher unbeobachteter Heterogenität wenden wir für die Panel-Teilnehmer Latent-Class-Modelle an (Wedel und Kamakura, 2001; Jedidi, Krider und Weinberg, 1998), bei denen die jeweiligen Objekte nicht deterministisch einem Segment zugeordnet werden, sondern mit einer bestimmten Wahrscheinlichkeit einem Segment zugehörig sind.¹⁶ Geschätzt wird in diesem Fall das Basis-Modell mit einem Mixture-Regression-Modell auf Grundlage der Normalverteilungsannahme (Wedel und Kamakura, 2001).¹⁷ Hierbei wird simultan eine Segmentlösung gesucht und je Segment eine Regressionsgleichung geschätzt. Anhaltspunkte zur Bestimmung der Anzahl der Segmente liefern die Informationskriterien.¹⁸ Die Informationskriterien basieren auf dem Log-Likelihood, beinhalten aber eine unterschiedliche Betrachtung des „Strafterms“ für die zusätzlich zu schätzenden Parameter für jedes neue Segment (Wedel und Kamakura, 2001). Tabelle 7 zeigt die jeweiligen Kriterien für mehrere Segmente. Die minimalen Werte für die Informationskriterien (z. B. CAIC oder BIC; Ramaswamy et al., 1993) finden sich bei der Drei-Segment-Lösung.

Auf der Basis der Informationskriterien ist die Drei-Segment-Lösung zu präferieren (knapp bessere Werte als die Zwei-Segment-Lösung). Allerdings haben alle identifizierten Drei-Segment-Lösungen das Problem, dass das ursprüngliche Segment 2 mit den hohen

Tab. 7. Gütekriterien der Segmentlösung (n=440)

Kriterium	Ein Segment	Zwei Segmente	Drei Segmente
Log-Likelihood	-3756,5	-3500,7	-3440,6
AIC	7535,0	7047,4	6951,2
CAIC	7591,0	7164,4	7129,2
MAIC	7546,0	7070,4	6986,2
BIC	7580,0	7141,4	7094,2
Entropie	1	0,67	0,70
Freiheitsgrade	11	23	35

Kosten in zwei neue Segmente aufgesplittet wird, wobei das neue Segment 3 sehr klein ist (2% der Stichprobe – also etwa 9 Personen mit sehr hohen Kosten). Die geschätzten Parameter für das kleine Segment weisen zudem sehr hohe t-Werte auf, so dass möglicherweise ein Identifikationsproblem besteht. Dieses Identifikationsproblem entsteht insbesondere dann, wenn sehr kleine Segmente mit wenigen Beobachtungen vorliegen (Wedel und Kamakura 2001, S. 109). Aufgrund der inhaltlich besseren Struktur und dem offensichtlichen Identifikationsproblem bei mehr als zwei Klassen wählen wir die Zwei-Segment-Lösung, die zwar etwas schlechtere Informationskriterien aufweist, aber die stabilere Lösung darstellt und inhaltlich plausibel ist.

Die Entropiestatistik, die ein verdichtetes Maß darstellt, inwieweit sich die Objekte den jeweiligen Segmenten eindeutig zuordnen lassen, weist für die Zwei-Segmentlösung einen zufrieden stellenden Wert von 0,67 aus (Ramaswamy et al., 1993).

In Tabelle 8 werden die geschätzten Parameter dargestellt. In Spalte 1 werden die OLS-Schätzergebnisse für das gesamte Sample der ersten Welle berichtet. Spalte 2 weist im Vergleich hierzu die OLS-Schätzparameter für die Panelstichprobe aus. Es zeigt sich, dass sich die signifikanten Parameter zwischen den beiden Samples nur gering unterscheiden. Der Schätzparameter für die Variable „Vorbeugung“ ist in der Welle 1 nun zum 5%-Niveau signifikant. Umgekehrt ist der t-Wert der Variable „Verbreitung an Unbekannte“ im reduzierten Sample geringer, so dass das 5% Signifikanzniveau nicht erreicht wird. Dennoch weisen beide Variablen einen t-Wert größer eins auf, so dass sie Prognoserelevanz besitzen (Haitowski, 1969; Hansen, 1987). Die OLS für die Welle 1 weist eine gute statistische Validität auf. So können 32% bzw. 34% der Varianz (R^2) durch die Modelle erklärt werden. Spalte 3 listet die Ergebnisse der OLS für Welle 2. Auch hier ist die Erklärungsgüte mit einem R^2 von 0,23 zufrieden stellend. Wie in Endnote 13 diskutiert, konnten nicht alle Variablen neu für die zweite Welle erhoben werden, so dass die Ergebnisse der OLS in Spalte 3 auf der Annahme basiert, dass die fünf Variablen *Überprüfung*, *Empfindlichkeit*, *Eigenschaften*, *Verbreitung an Bekannte* und *Unbekannte* konstant bleiben. Bevor auf die inhaltlichen Ergebnisse eingegangen wird, soll zunächst die Validität der Mixture-Regression-Lösung untersucht werden.

Wie in Tabelle 7 bzw. 8 verdeutlicht, werden zwei Klassen identifiziert, die sich in ihren Kosten erheblich unterscheiden. Im Segment 1 (S_1 , Spalte 4, Tabelle 8), welches mit einem Mischungsanteil von 66,6% das größere Segment darstellt, sind die Arbeitszeitverluste relativ gering (Mittelwert = 475 Minuten). Im Segment 2 (Spalte 6) sind die Kosten deutlich höher. So weist das mit 33,4% der Befragten besetzte Segment 2 (S_2) mittlere

Tab. 8. Vergleich der Modellparameter im Basismodell und in den Segmenten für die Teilnehmer des Panels (Segmentierungsbasis Welle 1)

	(1) OLS Welle 1 Regressions- parameter (t-Statistik)	(2) OLS Welle 1 (Panel) Regressions- parameter (t-Statistik)	(3) OLS Welle 2 (Panel) Regressions- parameter (t-Statistik)	(4) Segment 1 (66,6%) Mittlere Kosten = 475,4 Min. Regressions- parameter (t-Statistik)	(5) Mittelwert- veränderung von Welle 1 zu 2 Mittelw. Welle 1 (Mittelw. Welle 2)	(6) Segment 2 (33,4%) Hohe Kosten = 3.302,2 Min. Regressions- parameter (t-Statistik)	(7) Mittelwert- veränderung von Welle 1 zu 2 Mittelw. Welle 1 (Mittelw. Welle 2)
Spam-Anzahl	20,97 (11,93)	15,51 (6,46)	14,81 (5,67)	2,04 (2,07)	9,92 (10,86)	14,74 (3,53)	40,21 (31,85)
Vorbereitung	-94,80 (-1,87)	-162,59 (-2,30)	-3,44 (-0,05)	-1,59 (-0,08)	2,25 (2,38)	-176,80 (-0,92)	2,11 (2,34)
Überprüfung	165,43 (2,03)	263,64 (2,50)	139,57 (1,37)	42,43 (1,44)	3,67 (keine Messung)	465,72 (1,43)	3,73 (keine Messung)
Verhaltens- änderung	465,83 (7,98)	492,62 (5,97)	457,29 (5,48)	106,14 (4,06)	1,86 (1,89)	543,51 (1,80)	2,86 (2,38)
Empfindlich- keit	3,13 (0,04)	-20,32 (-0,21)	-59,44 (-0,65)	24,30 (0,89)	3,23 (keine Messung)	-133,84 (-0,61)	3,24 (keine Messung)
Eigenschaften	266,41 (4,34)	233,06 (2,83)	128,62 (1,68)	-13,37 (-0,67)	3,88 (keine Messung)	536,06 (1,23)	4,49 (keine Messung)
Informations- stand	-24,56 (-0,58)	21,92 (0,36)	-82,98 (-1,33)	7,04 (0,48)	2,61 (3,07)	67,24 (0,37)	2,84 (3,14)
Verbreitung Bekanntheit	-27,73 (-0,37)	7,23 (0,07)	-38,56 (-0,39)	-36,86 (-1,28)	1,48 (keine Messung)	157,63 (0,51)	1,41 (keine Messung)
Verbreitung Unbekanntheit	115,33 (2,35)	106,64 (1,61)	43,20 (0,68)	73,13 (4,05)	1,79 (keine Messung)	50,39 (0,20)	2,21 (keine Messung)
Konstante	-1678,76 (-3,62)	-1866,12 (-3,13)	-690,95 (-1,17)	-7,22 (-0,04)	-	-2975,30 (-1,15)	-
N	1.000	440	440	325	325	115	115
R ²	0,32	0,34	0,23			0,45	

Für die Regressionsergebnisse (Spalten 1–4 und Spalte 6) bedeuten fett gedruckte Werte, dass der entsprechende Regressionsparameter signifikant ($p < 0,05$) ist. Bei den Mittelwertvergleichen in den Spalten 5 und 6 bedeutet ein fett gedruckter Wert, dass sich die Mittelwerte signifikant ($p < 0,05$) unterscheiden (t-Test bei gepaarten Stichproben).

Arbeitszeitverluste von 3.302 Minuten aus. Die Zwei-Segment-Lösung weist ein R^2 von 0,45 aus und zeigt auf, dass im Vergleich zur Ein-Segment-Lösung (Spalte 2) deutlich differenziertere Ergebnisse erzielt werden. Während die Anzahl der Spam-Mails und die – seitens der Mitarbeiter – wahrgenommene Notwendigkeit zur Verhaltensanpassung über alle Segmente hinweg zu steigenden Kosten führen, ist die Wirkung bei anderen Variablen heterogener. Die Spalten 5 und 7 stellen den Vergleich der Mittelwerte der Variablen von Welle 1 und 2 in den Segmenten dar.

Es lassen sich folgende inhaltliche Ergebnisse festhalten. Die Anzahl der **Spam-Mails** hat plausiblerweise einen signifikanten Einfluss auf die Kosten in beiden Wellen und in beiden Segmenten der ersten Welle. Die Betrachtung der Segmentlösungen zeigt ein genaueres Bild auf: Im Segment 1, dem Segment mit den geringen Arbeitszeitverlusten, ist der Schätzparameter deutlich kleiner (2,04) als im Segment 2 (14,74). Es zeigt sich zudem (Spalten 5 bzw. 7), dass sich die mittlere Anzahl der Spam-Mails in den beiden Segmenten gegensätzlich über die Zeit entwickelt: Im S_1 steigt der Mittelwert von 9,92 (Welle 1) auf 10,86 (Welle 2) Spam-Mails pro Tag substantiell an. Im S_2 kann hingegen ein signifikantes Absinken von 40,21 auf 31,85 Spam-Mails pro Tag beobachtet werden. Demnach kann festgehalten werden, dass die Anzahl der Spam-Mails bei denjenigen abnimmt, die bereits hohe Kosten hatten, während die Menge bei den verbleibenden Nutzern weiter zunimmt. Dieser Effekt spiegelt sich auch in der aggregierten Betrachtung wider (Spalten 2 und 3), denn der Regressionsparameter sinkt von 15,51 in Welle 1 auf 14,81 in Welle 2. Dennoch bleibt die Spam-Menge auch in Welle 2 der zentrale Einflussfaktor auf die Arbeitszeitverluste (gemessen über standardisierte Regressionskoeffizienten).

Vorbeugende Maßnahmen seitens des Mitarbeiters reduzieren die Arbeitszeitverluste nur zum Teil. Während in der aggregierten Betrachtung in Welle 1 dieser Faktor einen signifikant negativen Einfluss aufweist (-162,59), lässt sich dieser Einfluss nicht in den Segmenten finden. Zwar ist das Vorzeichen des Schätzkoeffizienten in beiden Segmenten negativ, aber die jeweiligen t-Werte sind in beiden Segmenten kleiner als Eins, so dass von keinem substantiellen Beitrag dieser Variablen ausgegangen werden kann. Die Betrachtung der jeweiligen Segmentmittelwerte zeigt, dass sich kaum eine Veränderung über die Zeit eingestellt hat.

Umfangreiche Maßnahmen zur **Überprüfung** der E-Mails hinsichtlich der Inhalte weisen einen kostensteigernden Effekt auf. Die Wirkung dieser Variablen ist in den aggregierten Analysen der Welle 1 jeweils hochsignifikant, ein Einfluss auf die Kosten der zweiten Welle kann jedoch nicht festgestellt werden. Wir nehmen an, dass sich diese Variable nicht über die Zeit ändert. Dementsprechend wurde diese Variable in der zweiten Welle nicht erneut gemessen.¹³ In den beiden Segmenten wird deutlich, dass die Wirkung im Vergleich zu der aggregierten Schätzung nur moderat ist. Interessanterweise zeigt sich auch kein bedeutsamer Mittelwertunterschied zwischen den beiden Segmenten – die Werte liegen mit 3,67 für S_1 und 3,73 für S_2 nahezu auf gleichem Niveau.

Einen signifikanten und substantiellen Einfluss über alle Gruppen finden wir bei der Notwendigkeit der **Verhaltensänderung** durch Spam. Einige Befragte reduzieren die E-Mail-Nutzung aufgrund von Spam und verlieren ihr Vertrauen in das Kommunikationsmittel. Dieser eher auf psychologische Reaktanz zurückzuführende Aspekt beeinflusst die Kosten. Der Grund ist jedem Spam-geplagten Nutzer ersichtlich, der einmal über ein Web-Interface seinen E-Mail-Account abgerufen und dabei nicht seine normale E-Mail-

Software (z.B. Outlook) verwendet hat, die ggf. die Spam-Mails herausfiltert. Das manuelle Filtern der E-Mails bei langsamen Internetverbindungen (z. B. im Hotel) wird dann zur Qual – und führt zu Arbeitszeitverlusten beim Mitarbeiter. Es zeigt sich, dass der Parameter in der aggregierten OLS-Schätzung von 492,62 (Welle 1) auf 457,29 (Welle 2) sinkt. Dieses Absinken lässt sich vor allem durch das Segment 2 erklären, denn dort sinkt der Mittelwert von 2,86 (Welle 1) auf 2,38 (Welle 2).

Interessanterweise finden wir keinen Einfluss der **Spam-Empfindlichkeit** auf die verlorene Arbeitszeit. So macht es keinen Unterschied, ob jemand Witz- oder Werbe-Mails als Spam ansieht oder nicht – die Kosten werden davon nicht signifikant beeinflusst. Damit ist Spam nicht ein Problem von empfindlichen Mitarbeitern, sondern betrifft definitionsunabhängig alle Mitarbeiter. Dieser Befund ist unerwartet, denn wie in Tabelle 6 dargestellt ist, zeigt sich durchaus eine substantielle Standardabweichung von 0,623 um den Mittelwert von 3,256 für diese Variable, so dass die Empfindlichkeit über die Arbeitnehmer variiert.

Signifikant hingegen ist die Sorge vor bestimmten **Eigenschaften** von Spam, die einen kostensteigernden Effekt in den aggregierten Analysen vor allem in Welle 1 aufweist. In der zweiten Welle ist dieser Effekt zwar nicht mehr hochsignifikant, aber mit einem t-Wert von 1,68 durchaus substantiell. Die segmentspezifische Betrachtung offenbart, dass die Kosten vor allem durch die Wahrnehmungen des zweiten Segments determiniert werden. So gilt demnach, dass mit zunehmend negativer Wahrnehmung der Eigenschaften von Spam (z. B. Viren) die Kosteneffekte steigen. Dieser Effekt ist unmittelbar einsichtig und unterstreicht, dass die Mitarbeiter über Spam und den Auswirkungen eines nicht fachgerechten Handlings (z. B. durch das Öffnen von ausführenden Dateien, die möglicherweise Viren beinhalten) informiert werden sollten. Zwar führt dies zu Kostensteigerungen, aber es können möglicherweise größere Kosten (durch Virenbefall und Datenverlust) vermieden werden.

Interessanterweise liefert die Variable, die den **Informationsstand** zu Spam misst, keinen signifikanten Einfluss auf die Kosten. Schließlich hat die Verbreitung der eigenen E-Mail-Adresse nur dann einen signifikanten und kostensteigernden Einfluss, wenn sie an Unbekannte weitergegeben wird. Der Einfluss manifestiert sich vor allem in den signifikant steigenden Kosten im Segment 1.

Zur Validierung der Ergebnisse wurden alle hier diskutierten Einflussfaktoren auch auf deren Wirkung auf die wahrgenommene Belästigung durch Spam regressionsanalytisch betrachtet. Die abhängige Variable wurde auf einer 5-Punkt-Likert-Skala gemessen, so dass der Einsatz von Ordered Regressions erforderlich scheint (Tabelle 4). Für die Schätzung der Regression greifen wir auf das Schwellenwertmodell zurück (McCullagh, 1980). Der Literatur folgend benutzen wir die Logit-Funktion als Verknüpfungsfunktion (McCullagh und Nelder, 1989). Die Modellgüte lässt sich anhand der R²-Werte nach Nagelkerke (Welle 1: 0,46 und Welle 2: 0,37) bzw. McFadden (0,19 und 0,14) mit gut bezeichnen.¹⁹ Wir haben die Ergebnisse der Ordered Regression mit einer OLS verglichen und kommen zu nahezu identischen Werten der Regressionskoeffizienten. Des Weiteren wurde auf der Basis des OLS-Modells geprüft, ob auch hier Latent-Class-Verfahren notwendig sind, um die unbeobachtete Heterogenität zu kontrollieren. Die Informationskriterien weisen aber auf eine Ein-Segment-Lösung hin. Die Ordered Regression der wahrgenommenen Belästigung bestätigt im Wesentlichen die Ergebnisse der aggregierten OLS-Schätzungen für die Arbeitszeitverluste in Tabelle 8. Im Vergleich zur OLS nimmt jedoch in Welle 1 der signifikante Einfluss der Vorbeugung und auch der Überprüfung ab. In Welle

2 verstärken sich die Effekte der Spam-Eigenschaften und des Informationsstands über Spam. Dass diese einen größeren Einfluss auf die Störung als auf die Kosten haben, ist plausibel. Da beide Variablen in der OLS auch schon mit einem t-Wert von größer 1 ausgewiesen wurden, sind die Anpassungen nur von marginaler Relevanz. Deutlicher wirkt in Welle 2 die Weitergabe der E-Mail-Adresse an Unbekannte auf die Belästigung ein. Obgleich einige Effekte aufgrund der veränderten Messung der Kosten verstärkt wirken, werden die prinzipiellen Ergebnisse auch bei einer anderen Messung der Kosten gestützt.

IV. Analyse des Einflusses von Filtermaßnahmen auf die Kosten

Die Wirkung eines Spam-Filters auf die Arbeitszeitverluste bzw. auf die wahrgenommene Belästigung ist endogen, was sich anhand eines Mittelwerttests verdeutlichen lässt: Abbildung 7 listet sowohl die Mittelwerte für (1) die Arbeitszeitverluste (inkl. der eingesetzten Zeit zur Installation des Filters), (2) die wahrgenommene Belästigung durch Spam, (3) die Kosten durch mögliche Fehlklassifikationen des Filters und schließlich (4) die Arbeitszeitverluste durch die Bearbeitung von Spam ohne Berücksichtigung der Installation einer Software.

Es zeigt sich, dass die Kosten bei den 227 Nutzern mit Spam-Filtern (Welle 1) gegenüber den 213 Mitarbeitern ohne Filter signifikant höher sind. Es bleibt jedoch unklar, ob die Nutzer aufgrund der Spam-Belastung und den daraus resultierenden Kosten den

Welle 1		Filter (n=227)		Kein Filter (n=213)	
Kosten		1665,29		733,57	
Störung		3,86		2,94	
Fehlklassifikation		11,17		9,93	
Spamzeit p.a. ¹		1607,93		733,57	
Kosten		1730,96	1597,86	968,31	616,20
Störung		3,97	3,75	3,23	2,80
Fehlklassifikation		11,09	11,25	10,21	9,80
Spamzeit p.a. ¹		1668,48	1545,76	968,31	616,20
▼					
Welle 2		Schutz verstärkt n=115	Schutz konstant n=112	Filter eingerichtet n=71	Kein Filter n=142
Kosten		1436,87	1305,32	976,47	655,80
Störung		3,49	3,54	3,18	2,78
Fehlklassifikation		11,89	12,54	10,48	10,78
Spamzeit p.a. ¹		1375	1289,06	941,90	655,80
▼					
Ersparnis		Ersparnis im Vergleich zu Welle 1			
Kosten		294,08*	292,54*	-8,16 ^{ns}	-39,61 ^{ns}
Störung		0,48***	0,21*	0,04 ^{ns}	0,01 ^{ns}
Fehlklassifikation		-0,80*	-1,29**	-0,27 ^{ns}	-0,99**
Spamzeit p.a. ¹		293,48*	256,70*	26,41 ^{ns}	-39,61 ^{ns}

¹ Spam-Bearbeitungszeit pro Tag * 250 Tage
 * Signifikant mit p<0,1; ** Signifikant mit p<0,05; *** Signifikant mit p<0,01 (Signifikanzniveaus beziehen sich auf einen t-Test, der die Mittelwerte („within subjects“) zwischen Erhebungzeitpunkt 1 und 2 vergleicht.)

Abb. 7. Kostenwirkungen von Spam-Filtern bei den Teilnehmern des Panels (n=440)

Spam-Filter installierten, oder ob der Filter die Kosten steigen lässt. Zwar kann eher ersteres vermutet werden (dies kann durch die Gruppe, die in der ersten Welle keinen Filter installiert hatte, belegt werden), aber die Vermutung kann mit den vorliegenden Querschnittsdaten der ersten Welle nicht für alle Nutzer gestützt werden. Es zeigt sich aber im Vergleich zwischen der ersten und zweiten Welle, dass die Mitarbeiter, die bereits zum Zeitpunkt der ersten Befragung einen Filter installiert hatten, zum Zeitpunkt der zweiten Befragung deutlich niedrigere Kosten aufwiesen. So sinkt die verlorene Arbeitszeit von 1.665 in Welle 1 auf 1.436 Minuten (für diejenigen, die den Schutz verstärkten) bzw. auf 1.305 Minuten (für diejenigen, die den Filter so beließen) in Welle 2. Dies deutet darauf hin, dass zum einen eher die Personen, die hohe Kosten haben, einen Spam-Filter installieren. Zum anderen scheint der Filter aber längerfristig die Kosten zu senken. Die Kostenersparnisse (within subjects) sind in der Zeile „Ersparnis“ in Abbildung 7 aufgezeigt. Es zeigt sich, dass die Nutzer, die bereits in Welle 1 einen Filter installiert haben, signifikante Einsparungen hinsichtlich der verlorenen Arbeitszeit und der Belästigung erzielen können. Nur steigt dort auch die Sorge vor Fehlklassifikationen signifikant. Diese Sorge spiegelt sich allerdings nicht in einem negativen Nettoeffekt der Arbeitszeiterparnis durch verstärkte Kontrolle wider, so dass der Einsatz von Filtern langfristig positive Effekte mit sich bringt. Als Ursache hierfür können vor allem Lerneffekte vermutet werden, die jedoch nicht die Sorge vor Fehlklassifikationen verringern.²⁰ Bei Nutzern mit niedrigen Kosten scheint eine Installation des Filters kurzfristig jedoch keine Kostenreduktion zu bewirken.

F. Schlussfolgerungen

Die Ergebnisse dieser Untersuchung zeigen auf, dass die Kosten auf Seiten des Providers im Vergleich zu den Kosten durch Arbeitszeitverluste auf Mitarbeitererebene zu vernachlässigen sind. So stehen den Kosten des Rechenzentrums in Höhe von knapp 15.000 € Kosten auf Mitarbeitererebene gegenüber, die um ein Vielfaches höher sind. Unter der Annahme, dass die Kosten repräsentativ für die Universität erhoben wurden, ist von Gesamtkosten aufgrund von Arbeitszeitverlusten von ca. 2,2 Mio. Euro pro Jahr auszugehen (Basis: 5.000 Mitarbeiter).

Als die zentralen Kostentreiber auf Mitarbeitererebene sind hierbei regressionsanalytisch die Anzahl der empfangenen Spam-Mails und die durch Spam-Mails implizierte Verhaltensänderung der Mitarbeiter identifiziert worden. Diese Variablen haben einen durchgängig signifikanten Einfluss auf die abhängigen Kostenvariablen. Mitarbeiter, die ihre E-Mail-Adresse umfangreich an Unbekannte verbreitet haben, weisen signifikant höhere Arbeitszeitverluste auf. Zudem steigen die Kosten mit dem Umfang des Überprüfens von Spam-Mails hinsichtlich ihres Inhalts. Wir finden weder einen signifikanten Einfluss dahingehend, dass ein besserer Informationsstand der Personen einen Einfluss auf die Kosten hat, noch übt die Vorbeugung und die individuelle Spam-Empfindlichkeit einen erkennbaren Einfluss aus. Die Installation eines Spam-Filters führt zu Arbeitszeitverlusten, so dass die Kosten nicht kurzfristig, sondern eher langfristig gesenkt werden können.

Die Ergebnisse unserer Studie ermöglichen uns das Bereitstellen von Managementimplikationen. Die Analyse zeigt, dass Spam-belastete Organisationen einen zentralen Schutz installieren sollten, der Spam-Mails über Blacklisten bereits zentral abblockt.

Somit sinkt die Anzahl der zugestellten Spam-Mails an die Mitarbeiter. Zudem eignet sich eine zentrale Lösung zur Markierung der Mails hinsichtlich der Spam-Wahrscheinlichkeit – ohne diese zu filtern. Die niedrigen zentralen Kosten im Rechenzentrum deuten darauf hin, dass die erste Barriere gegen Spam am effizientesten zentral umgesetzt werden kann. Außerdem sollte diese Maßnahme um die Möglichkeit ergänzt werden, auf individueller Ebene in den Prozess der Spam-Klassifikation einzugreifen. Damit wird die Heterogenität in den Präferenzen und Wahrnehmungen der Nutzer berücksichtigt und somit die Sorge vor Fehlklassifikationen reduziert. Dieser integrative Ansatz dürfte dem Problem am ehesten gerecht werden (Park und Deshpande, 2006). Außerdem kann der Analyse entnommen werden, dass ein Spam-Filter erst ab einem gewissen individuellen Spam-Volumen zu einer Kostenreduktion führt. Das Ergreifen von individuellen Maßnahmen sollte daher nicht pauschal, sondern vor allem bei belasteten Mitarbeitern vorgenommen werden. Die Ergebnisse zeigen weiterhin, dass sich der Informationsstand zum Thema Spam auf niedrigem Niveau verbessert hat. Daher sollten Unternehmen den Wissensstand ihrer Mitarbeiter in zwei Richtungen erweitern. Zum einen sollten eher Möglichkeiten zur effizienten Spam-Abwehr als zur Vorbeugung vermittelt werden, da letztere keinen großen Einfluss auf die Kosten zu haben scheint. Ein verbesserter Informationsstand dürfte zum anderen auch dazu führen, dass der Einfluss der Spam-Eigenschaften (d. h. Gefährdung der Computer-Sicherheit) auf die Kosten zurückgeht. Eine Verbesserung der wahrgenommenen individuellen Kompetenz dürfte den Anwendern somit das Gefühl geben, das Spam-Problem besser im Griff zu haben und somit kostenreduzierend wirken.

Die Ergebnisse dieser Studie bieten auch Implikationen für Unternehmen, die über das Internet Direkt-Marketing-Maßnahmen vornehmen möchten. Um weiterhin effizient mit dem Kunden bspw. über Newsletter zu interagieren, ist es von äußerster Wichtigkeit, dass die Newsletter nicht als Spam identifiziert werden. Die erste Hürde, die ein Newsletter dafür nehmen muss, ist der zentral installierte Spam-Filter (bspw. SpamAssassin). In diesem Kontext dominieren technische Aspekte – es sollten z.B. keine Elemente einer E-Mail manipuliert werden und es sollte nicht versucht werden, die Absenderadresse auf irgendeine Art zu verstecken. Darüber hinaus gilt es sicherzustellen, dass die angesprochenen Personen den Newsletter nicht als Spam wahrnehmen. Hier nimmt die Formulierung der Betreffzeile eine zentrale Position ein, da diese – wie wir oben zeigen konnten – von den meisten Befragten zu Beantwortung der Frage, ob sich um Spam handelt, herangezogen wird. Sollte ein Newsletter als Spam wahrgenommen werden, kann es zu Reaktanz kommen, die die Marke schädigen kann. Dieser Effekt kann auch eintreten, wenn eine Möglichkeit fehlt, den Newsletter abzubestellen, denn von den meisten Nutzern wird Spam insbesondere deshalb als kritisch wahrgenommen, weil deren Verbreitung nicht zu stoppen ist. Die ermittelten Ergebnisse legen des Weiteren nahe, dass Newsletter nur dann eingesetzt werden sollten, wenn sie dem Kunden substanziellen Nutzen stiften. Das schließt ein, dass sie nicht in einer zu häufigen Frequenz und unter keinen Umständen unaufgefordert zugesandt werden. Das E-Mail-Marketing muss also im Kontext des Permission-Marketings betrieben werden (Drèze, 2005). Verhalten, das von den E-Mail-Empfängern als unlauter wahrgenommen wird, kann dazu führen, dass das versendende Unternehmen auf individuellen oder zentralen Blacklists geführt wird, was den zukünftigen Empfang von Newslettern des entsprechenden Absenders sehr unwahrscheinlich werden lässt. Die ausgeprägte Heterogenität in der individuellen Spambelastung stellt Un-

ternehmen außerdem vor die Herausforderung, die jeweiligen Segmente mit der richtigen Intensität des Direkt-Marketing per E-Mail anzusprechen. Die bisher häufig zu beobachtende Praxis, alle Kunden in der Hinsicht gleich zu behandeln, wird der in dieser Untersuchung festgestellten Heterogenität der Spam-spezifischen Präferenzen nicht gerecht.

Jede empirische Studie weist Potenzial für anknüpfende Arbeiten auf. Zum einen wäre es lohnenswert, den Panel-Ansatz weiter zu nutzen, um die kostenreduzierenden Wirkungen eines Spam-Filters langfristig zu untersuchen. Zum anderen ist zu prüfen, ob die Ergebnisse auch für andere Organisationen gelten, um so generalisierbare Resultate zu erhalten (auch international), oder ob systematische Unterschiede zwischen verschiedenen Organisationsformen bestehen. Ein weites Forschungsfeld eröffnet sich auch dem Marketing – denn die Optimierung des Direkt-Marketings im Zeitalter von Spam-Filtern stellt eine große Herausforderung dar. So sollten bisherige Optimierungsmodelle um die Komponente der Empfängerakzeptanz erweitert werden, wobei die Subjektivität der Wahrnehmung sowie die begrenzte Informationsverarbeitungskapazität der Empfänger berücksichtigt werden sollten. Schließlich sind auch Arbeiten von Interesse, die eher dem volkswirtschaftlichen Bereich zuzurechnen sind, denn letztlich können Marktmechanismen implementiert werden, die die Kosten der Kommunikation auf den Absender wälzen und so das bisherige Geschäftsmodell der Spammer ad absurdum führen. Zwar kommt auch per normaler Post eine Flut von Werbung ins Haus, aber letztlich muss der Absender hierfür Porto entrichten. Dies gilt im Internet noch nicht – und daher lohnt sich Spam.

Anmerkungen

* Die Autoren bedanken sich sehr herzlich bei den Mitarbeitern des Rechenzentrums der Christian-Albrechts-Universität zu Kiel für die intensive Kooperation. Wir danken zudem Prof. Dr. Dr. h.c. Sönke Albers, Christian Barrot, Eva Blömeke und Christina Schmidt-Stölting für die zahlreichen Hinweise zu früheren Fassungen des Aufsatzes. Auch die beiden Gutachter und das Herausgeberteam der ZfB haben zahlreiche Verbesserungsvorschläge gemacht, die zu einer substantiellen Verbesserung des Manuskripts geführt haben.

- 1 Eigentlich bezeichnet der Name „Spam“ Dosenfleisch der amerikanischen Firma HORMEL FOODS (Shoulder of Pork And ham). Die Nutzung des Wortes für Werbemails ist höchstwahrscheinlich auf einen Sketch der englischen Komikergruppe MONTY PYTHON zurückzuführen, in dem der Name dieses Produkts innerhalb weniger Minuten knapp einhundert Mal fällt und damit jede Konversation übertönt.
- 2 Die im Vergleich zur globalen Spam-Belastung höhere Spam-Quote an der Universität hängt damit zusammen, dass viele Universitätsmitarbeiter schon sehr frühzeitig mit den E-Mail-Adressen auf den Websites der Universität gelistet waren, so dass die Spammer bereits zu Beginn des Internetaufkommens auf diese Adressen zugreifen konnten.
- 3 Zentral installierte Spam-Filter blocken auf Basis klar definierter Regeln als Spam identifizierbare E-Mails ab bzw. markieren sie vor dem Weiterleiten an den Endnutzer als Spam (z. B. SpamAssassin). Dezentral arbeitende Filter überprüfen die E-Mail erst nach dem Abruf auf ihre Spam-Eigenschaften und sortieren verdächtige E-Mails aus (z. B. Microsoft Outlook). Zahlreiche Varianten und Kombinationen werden in der Praxis eingesetzt, wobei die meisten Systeme individuell angepasst werden können und lernfähig sind (Bayes-Filter; Schryen, 2004; Park und Deshpande, 2006).
- 4 Auch eine a posteriori Betrachtung der Validität mittels des Tetrad-Tests eignet sich nur für die Prüfung, ob ein reflektives Messmodell nicht eher formativ gebildet werden sollte – jedoch nicht umgekehrt (Bollen und Ting, 2000; Gudergan, 2005; Weise, Salomo und Gemünden, 2006). Da der Tetrad-Test nicht in der Lage ist, ein theoretisch angebrachtes formatives Messmodell mit hoch korrelierenden Indikatoren von einem reflektiven Messmodell zu unterscheiden, kann auch der Tetrad-Test eine korrekte formative Spezifikation des Messmodells nicht mit hinreichender Sicherheit bestätigen bzw. widerlegen. Das bedeutet, dass der Test für die Prüfung reflektiver Messmodelle geeignet ist, um so eine Fehlspezifikation zu vermeiden, jedoch die Validität formativer Skalen nur aufgrund inhaltlicher Überlegungen und nicht anhand von statistischen Maß-

- zahlen bestimmt werden kann. Darüber hinaus wird in der Literatur darauf verwiesen, dass die Validität des Messmodells anhand eines theoretisch begründeten und empirisch belegten Zusammenhangs zu anderen latenten Konstrukten des Modells mittels eines Two-Construct Modells überprüft werden kann. Aber auch dies ist nicht unkritisch zu sehen (Krafft, Götz und Liehr-Gobbers, 2005). Wird die Übereinstimmung eines formativen Konstrukts mit einer latenten reflektiven Variablen überprüft, so lassen sich auch in diesem Fall die klassischen Kritikpunkte in Bezug auf die reflektiven Konstrukte erneuern. Wird die Validität anhand des Zusammenhangs mit anderen formativen Konstrukten bestimmt, so ist offensichtlich die erste formative Messung nicht ausreichend vorgenommen worden (Reinstrom 2007, S. 108). Vor diesem Hintergrund liefert auch diese Vorgehensweise keine eindeutigen Aussagen bezüglich der externen Validität formativer Konstrukte. Wir danken einem anonymen Gutachter für diesen Hinweis.
- 5 Mit Ausnahme laufzeitabhängiger Kosten haben sich die Kosten im Laufe der Zeit nicht erheblich geändert. Dies ergab ein Gespräch mit den Verantwortlichen im Rechenzentrum im Mai 2006.
 - 6 Die Tatsache, dass einige Institutionen technisch, nicht aber administrativ, an das Rechenzentrum angeschlossen sind, hat diesen Vorgang erheblich erschwert. Aufgrund von Mitbestimmungsrechten mussten sich Personalräte, Rektorate etc. verschiedener Einrichtungen mit diesem Thema auseinandersetzen. Wegen der Schwierigkeiten der Datenerhebung in dieser zersplitterten Struktur wurden bei der Bestimmung der Kosten ausschließlich Zeitaufwendungen von Mitarbeitern des Rechenzentrums berücksichtigt (d.h. die wahren Kosten werden etwas unterschätzt).
 - 7 Dabei wurde vom durchschnittlichen Einkommen in dieser Gehaltsklasse (inklusive Lohnnebenkosten) von 4.500 EUR und einer Wochenarbeitszeit von 38,5 Stunden ausgegangen, so dass wir grob mit 30 EUR pro Stunde kalkulieren.
 - 8 Dies bezieht sich vor allem auf das Klinikum der Universität, so dass ein Bias in dieser Hinsicht besteht. Dieser ist jedoch wegen der besonderen Beschaffenheit der Arbeitsbedingungen in Kliniken nicht unbedingt als nachteilig für die Generalisierung der Ergebnisse für Unternehmen anzusehen.
 - 9 Die deskriptiven Analysen der Mitarbeitermerkmale „Alter“ und „Status“ ergaben nach Auskunft des Personaldezernats keine systematischen Abweichungen der Stichprobe von der gesamten Mitarbeiterstruktur der Universität. Allerdings bedeutet dies auch nicht zwangsläufig, dass eine in allen Belangen repräsentative Stichprobe vorliegt, denn es können weitere unbeobachtete Faktoren vorliegen, die einen möglichen Bias mit sich bringen.
 - 10 Dies gilt auch für die Wochenenden. Allerdings nutzen wir im Weiteren zur Berechnung der Arbeitszeitverluste nur die Anzahl der Arbeitstage (250 Tage) und nicht das gesamte Jahr (365 Tage).
 - 11 Hierbei nehmen wir an, dass die „verschwendete“ Zeit durch Spam tatsächlich einen Produktivitätsverlust darstellt und dass die verlorene Zeit nicht durch unbezahlte Überstunden oder angestiegener Produktivität ausgeglichen wird. Wir danken einem anonymen Gutachter für diesen Hinweis.
 - 12 Da wir dem Rektorat zugesagt hatten, den Fragebogen nicht in seiner vollen Länge erneut zu abzufragen und auch zusätzliche Fragen zur Zufriedenheit des Rechenzentrums gestellt haben, mussten wir auf die wiederholte Messung einiger Variablen verzichten. Es handelt sich um die Variablen der Indizes Spam-Überprüfung, Spam-Empfindlichkeit, Spam-Eigenschaften sowie die Verbreitung der E-Mail-Adresse an Bekannte und Unbekannte. Erneut messen wir vor allem die Variablen, von denen wir annehmen, dass sie sich im Vergleich zur ersten Welle geändert haben, weil die Befragten direkt ihr Verhalten ändern konnten (z. B. den Spam-Filter verschärfen) und so kurzfristige Kostenwirkungen entstehen können. Hierbei handelt es sich um die Spam-Menge, Spam-Vorbeugung und Verhaltensänderungen. Im Gegensatz hierzu stellen die anderen Variablen eher langfristig durch Lernen verstärkte Einstellungen dar oder sie lassen sich nicht rückgängig machen (z. B. die Verbreitung der E-Mail-Adresse), so dass wir annehmen, dass die einstellungsbezogenen Variablen eher konstant bleiben und daher auf die erneute Messung in der zweiten Welle verzichtet haben.
 - 13 Es könnte der Eindruck entstehen, dass wir die Anzahl der Spam-Mails zweifach berücksichtigen. Es handelt sich hierbei jedoch um unterschiedliche zu berücksichtigende Facetten: So messen wir bei den nutzerspezifischen Faktoren (Tabelle 5) die genaue Spam-Menge (in Anzahl Mails), während die wir bei den Spam-Eigenschaften mit der Variable w_4 (Tabelle 6) messen, inwieweit Personen Mails als Spam ansehen. Mit fünf Facetten erfassen wir, inwieweit es zu einer Einstellungsanpassung (von Mails zu Spam) kommt – nämlich, dass die Mails nicht angefordert wurden (w_1), die Mailflut trotz Gegenmaßnahmen nicht gestoppt werden kann (w_2), der Nutzer Angst vor dem Schadenspotenzial durch Viren etc. entwickelt hat (w_3), die Zeit, die mit Spam-Bearbeitung verbraucht wird (w_4) und schließlich mit (w_5) die wahrgenommene Menge an Spam-Mails, die zu einer Einstellungsänderung führen kann. Aufgrund der formativen Messung des Konstrukts ist es notwendig, diesen Aspekt mit einzubeziehen, um so die Messung des Konstrukts „Spam-Eigenschaften“ nicht zu verzerren. Zudem werden unterschiedliche Aspekte abgefragt, denn die wahrgenommene Spam-Belastung ist nicht gleichzusetzen mit der Anzahl der Spam-Mails, die empfangen wird. So werden 30

- Spam-Mails pro Tag von dem einen als nahezu unerträglich angesehen, während ein anderer sich über diese geringe Belastung freuen würde. Wir danken einem anonymen Gutachter für diesen Hinweis.
- 14 Die Korrelationstabellen können bei den Autoren direkt angefordert werden.
 - 15 Das Einbeziehen der Dummy-Variable „Filter“ in die folgenden Analysen führt dazu, dass das Vorzeichen des geschätzten Parameters positiv ist, d.h. dass die Kosten mit der Installation eines Filters steigen. Dies ist wahrscheinlich eine falsche kausale Beziehung, denn es ist wahrscheinlicher, dass eine Person mit hohen Kosten einen Spam-Filter nutzt. Dennoch würde das Einbeziehen der Variablen keine substantziellen Veränderungen der anderen Schätzparameter mit sich bringen – aus Gründen der Endogenität wird sie jedoch zunächst nicht weiter betrachtet.
 - 16 Die Ergebnisse der Mixture-Regression-Analyse für die 1.000 Teilnehmer der ersten Befragungswelle können direkt bei den Autoren angefordert werden.
 - 17 Die Schätzung der Modelle wird mit der Software Glimmix 3.0 vorgenommen (Wedel und Kamakura, 2001).
 - 18 Die Bestimmung der Anzahl der Segmente auf der Basis der Informationskriterien ist nicht unumstritten (siehe z.B. Celeux und Soromenho, 1996). Da aber die „wahre“ Anzahl und die Struktur der Segmente in der wirtschaftswissenschaftlichen Forschung zumeist unbekannt ist, können die Informationskriterien letztendlich nur als Indikatoren dienen – wir danken einem anonymen Gutachter für diesen Hinweis. Im Vergleich zu den häufig eingesetzten Clusteranalysen haben die Mixture-Regression-Modelle allerdings den Vorteil, dass sie mit den Informationskriterien zumindest Indikatoren liefern können, um so zu einer, wenngleich rein auf statistischen Kriterien basierenden, Segmentlösung zu kommen.
 - 19 Als Vergleichswert haben wir eine Standard-OLS gerechnet (R^2 Welle 1 = 0,43 und Welle 2 = 0,34).
 - 20 Die Korrelation zwischen Sorge vor Fehlklassifikation und Informationsstand in den jeweiligen Wellen liegt bei 0,071 bzw. 0,038 und ist nicht signifikant.

Literatur

- Ajzen, I. und M. Fishbein (1977): Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research, in: *Psychological Bulletin*, 84 (5), S. 888–918.
- Akcura, M.T. und K. Srinivasan (2005): Customer Intimacy and Cross-Selling Strategy, in: *Management Science*, 51 (5), S. 1007–1012.
- Albers, S. und L. Hildebrandt (2006): Methodische Probleme bei der Erfolgsfaktorenforschung – Messfehler, formative versus reflektive Indikatoren und die Wahl des Strukturgleichungs-Modells, in: *Zeitschrift für betriebswirtschaftliche Forschung*, 58 (1), S. 2–33.
- Armitage, C.J. und M. Conner (2001): Efficacy of the Theory of Planned Behaviour: A meta-analytic review, in: *British Journal of Social Psychology*, 40 (4), S. 471–499.
- Bollen, K.A. und K.-f. Ting (2000): A Tetrad Test for Causal Indicators, in: *Psychological Methods*, 5 (1), S. 3–22.
- BSA (2004): Verbraucher-Einstellung zu Spam in Deutschland, <http://www.bsa.org/germany/info/loader.cfm?url=/commonspot/security/getfile.cfm&pageid=20716&hitboxdone=yes>.
- Celeux, G. und G. Soromenho (1996): An entropy criterion for assessing the number of clusters in a mixture model, in: *Journal of Classification*, 13 (2), S. 195–212.
- Chu, W., E. Gerstner und J.D. Hess (1995): Costs and Benefits of Hard-Sell, in: *Journal of Marketing Research*, 32 (2), S. 97–102.
- Davis, F.D., R.P. Bagozzi und P.R. Warshaw (1989): User Acceptance of Computer Technology: A Comparison of two Theoretical Models, in: *Management Science*, 35 (8), S. 982–1003.
- Diamantopoulos, A. und H.M. Winklhofer (2001): Index Construction with Formative Indicators: An Alternative to Scale Development, in: *Journal of Marketing Research*, 38 (May), S. 269–277.
- Drèze, X. (2005): Lessons from the front line: two key ways in which the internet has changed marketing forever, in: *Applied Stochastic Models in Business and Industry*, 21, S. 443–448.
- Dwyer, F.R. (1997): Customer Lifetime Valuation to Support Marketing Decision Making, in: *Journal of Direct Marketing*, 11 (4), S. 6–13.
- Gudergan, S. (2005): PLS and Confirmatory Tetrad Testing for Formative Measurement Scales in Marketing, in: Paper presented at the 4th International Symposium on PLS and Related Methods, Barcelona.
- Haitowski, Y. (1969): A Note on the Maximization of R^2 , in: *American Statistician*, 23 (February), S. 20–21.
- Hansen, G. (1987): Multikollinearität und Prognosefehler, in: *Jahrbücher für Nationalökonomie und Statistik*, 203 (5-6), S. 517–531.

- Hesse, J., M. Krafft und K. Peters (2005): Grenzenloses Direktmarketing – Bestandsaufnahme, Trends und Ausblick, in: Krafft, M., et al. (Hrsg.): Internationales Direktmarketing, Wiesbaden, S. 1–43.
- Jarvis, C.B., S.B. Mackenzie und P.M. Podsakoff (2003): A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research, in: *Journal of Consumer Research*, 30 (September), S. 199–218.
- Jedidi, K., R.E. Krider und C.B. Weinberg (1998): Clustering at the Movies, in: *Marketing Letters*, 9 (4), S. 393–405.
- Joseph, K. und A. Thevaranjan (2005): Investigating Pricing Solutions to Combat Spam: Postage Stamp and Bonded Senders, in: Working Paper an der University of Kansas.
- Kommission der Europäischen Gemeinschaften (2004): Mitteilung der Kommission an das europäische Parlament, den Rat, den europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen über unerbetene Werbenachrichten (SPAM), http://europa.eu.int/information_society/topics/ecommerce/doc/useful_information/library/communic_reports/spam/spam_com_2004_28_de.pdf.
- Krafft, M. (2002): Kundenbindung und Kundenwert, Heidelberg.
- Krafft, M., O. Götz und K. Liehr-Gobbers (2005): Die Validierung von Strukturgleichungsmodellen mit Hilfe der Partial-Least-Squares (PLS)-Ansatzes, in: Bliemel, F., A. Eggert, G. Fassot und J. Henseler (Hrsg.), *Handbuch PLS-Pfadmodellierung*, Stuttgart, S. 71–86.
- Krafft, M. und K. Peters (2005): Empirical Findings and Recent Trends of Direct Mailing Optimization, in: *Marketing – Journal of Research and Management*, 1 (1), S. 26–40.
- Kraut, R.E., S. Sunder, R. Telang und J. Morris (2005): Pricing Electronic Mail to Solve the Problem of Spam, in: *Human-Computer Interaction*, 20, S. 195–223.
- Kuri, J. (2006): T-Online verzeichnet eine Milliarde Spam-Mails pro Tag, in: Heise Online, <http://www.heise.de/newsticker/meldung/72324>.
- Lwin, M.O. und J.D. Williams (2003): A Model Integrating the Multidimensional Developmental Theory of Privacy and Theory of Planned Behavior to Examine Fabrication of Information Online, in: *Marketing Letters*, 14 (4), S. 257–272.
- McCullagh, P. (1980): Regression Models for Ordinal Data, in: *Journal of the Royal Statistical Society*, 42 (2), S. 109–142.
- McCullagh, P. und J. Nelder (1989): *Generalized Linear Models*, London, Chapman Hall.
- McKnight, D.H., V. Choudhury und C. Kacmar (2002): Developing and Validating Trust Measures for e-Commerce: An Integrative Typology, in: *Information Systems Research*, 13 (3), S. 334–359.
- OECD (2005): Spam Issues in Developing Countries, <http://www.oecd.org/dataoecd/5/47/34935342.pdf>.
- o.V. (2004): Spam: The Serial ROI Killer, Nucleus Research.
- Park, J.S. und A. Deshpande (2006): Spam Detection: Increasing Accuracy with a Hybrid Solution, in: *Information Systems Management*, 23 (1), S. 57–67.
- Ramaswamy, V., W.S. DeSarbo, D. Reibstein und W.T. Robinson (1993): An Empirical Pooling Approach for Estimating Marketing Mix Elasticities with PIMS Data, in: *Marketing Science*, 12 (1), S. 103–124.
- Reinstrom, Christian (2007): *Steuerung des Markentransfererfolgs*, Wiesbaden.
- Rossiter, J.R. (2002): The C-OAR-SE procedure for scale development in marketing, in: *International Journal of Research in Marketing*, 19 (4), S. 305–335.
- Rossiter, J.R. (2005): Reminder: a horse is a horse, in: *International Journal of Research in Marketing*, 22 (1), S. 23–25.
- Schryen, G. (2004): Effektivität von Lösungsansätzen zur Bekämpfung von Spam, in: *Wirtschaftsinformatik*, 46 (4), S. 281–288.
- Schryen, G. (2006): A Formal Approach towards Assessing the Effectiveness of Anti-spam Procedures, in: *Proceedings of the 39th Hawai'i International Conference on System Sciences*.
- Sipior, J.C., B.T. Ward und P.G. Bonner (2004): Should Spam be on the Menu?, in: *Communications of the ACM*, 47 (6), S. 59–63.
- Topf, J. (2005): Ausgesiebt – Wie Mail-Provider gegen Spam vorgehen, in: *c't*, (11), S. 188–191.
- Vircom (2004): *Why Spammers SPAM*, White Paper at www.vircom.com.
- Wedel, M. und W.A. Kamakura (2001): *Market segmentation: conceptual and methodological foundations*, Boston, Mass.
- Weise, J., S. Salomo und H.G. Gemünden (2006): Reflexive oder formative Konstruktmessung: Tetradentest und Effekt auf Strukturgleichungsmodelle – eine empirische Annäherung., in: Working Paper an der Universität Graz.
- Zhang, L. (2005): The CAN-Spam Act: An Insufficient Response to the Growing Spam Problem, in: *Berkeley Technology Law Journal*, 20, S. 301–332.

Kosten und Kostentreiber von unerwünschten Werbemails (Spam) – Eine empirische Analyse auf Provider- und Anwenderseite

Zusammenfassung

In der Literatur wird diskutiert, dass unerwünschte Werbemails (Spam) sowohl auf individueller als auch auf volkswirtschaftlicher Ebene erhebliche Kosten verursachen, deren Höhe jedoch unbestimmt ist. Diese Fragestellung wird mit zwei empirischen Untersuchungen adressiert, in denen die Mitarbeiter einer Universität zu ihren Spam-Kosten befragt wurden. Die Daten werden mit Hilfe von Finite Mixture Regressions analysiert. In der Betrachtung der Kosten ergibt sich der Befund, dass die auf individueller Ebene entstehenden Kosten die zentralen Kosten auf IT-Ebene deutlich übersteigen. Es zeigt sich, dass die Spam-Belastung bei den Mitarbeitern heterogen verteilt ist und dass ein Spam-filter vor allem bei stark belasteten Nutzern zu einer Kostenersparnis führt. Um die Heterogenität in der Spam-Wahrnehmung und die Angst vor Fehlklassifikationen zu berücksichtigen, sollten die Nutzer die Möglichkeit haben, auf einfache Art auf den Filterprozess einzuwirken.

Costs and cost drivers of unsolicited bulk email (spam) – an empirical analysis

Summary

This paper empirically addresses the often claimed assumption that unsolicited bulk e-mail (spam) causes significant costs – on the individual level as well as for the entire society. To shed light on this problem, employees of a large university were interviewed twice in order to determine the individual costs caused by spam and to identify the relevant cost drivers using finite mixture regressions. The results indicate a heterogeneous distribution of spam load and the resulting costs. The panel approach yields the conclusion that spam filters lead to a reduction of costs only for users severely troubled by spam. Furthermore, people tend to become active against spam only when a certain threshold of spam burden is reached. The costs caused by spam for the users exceed the central IT-related expenses by far. For the implementation of spam filters the results suggest a two step procedure to be most efficient: central actions should be taken to block e-mails that can be unambiguously identified as spam. Further, users should be given the opportunity to influence the filtering process in order to account for heterogeneity and their fear of false classification.