

Marketing

Eine managementorientierte Einführung

von

Prof. Dr. Franz-Rudolf Esch, Prof. Dr. Andreas Herrmann, Prof. Dr. Henrik Sattler

4. Auflage

[Marketing – Esch / Herrmann / Sattler](#)

schnell und portofrei erhältlich bei beck-shop.de DIE FACHBUCHHANDLUNG

Thematische Gliederung:

[Marketing, Handelsmanagement](#) – [Marketing, Medien und Handel](#)

Verlag Franz Vahlen München 2013

Verlag Franz Vahlen im Internet:

www.vahlen.de

ISBN 978 3 8006 4691 3

$$\hat{p}(Y) = \frac{\exp(a + b \cdot X)}{(1 + \exp(a + b \cdot X))} = \frac{1}{(1 + \exp(-1 \cdot (a + b \cdot X)))} \quad (3)$$

Nach Transformation wird durch (3) die S-förmige logistische Funktion beschrieben, die auf das Intervall zwischen 0 und 1 definiert ist. Ein Beispiel für eine solche logistische Funktion ist in Abbildung C.28 aufgeführt.

Während bei der linearen Regression die Schätzer der unabhängigen Variablen den direkten Einfluss auf die abhängige Variable angeben, ist die Interpretation der Schätzer bei der logistischen Regression schwieriger. Sie geben den Einfluss auf die log odds an (Gleichung (2)). Dieser Wert sowie das weitere Vorgehen bei der logistischen Regression sollen an einem Beispiel erläutert werden.

Das Beispiel in Abbildung C.28 stellt eine Untersuchung der Teilnahme an fünf Wellen einer Online-Panelumfrage (Y, mit 0 = keine Teilnahme, 1 = Teilnahme) in Abhängigkeit von der Entlohnung der Panelmitglieder (E: Wert in Euro) sowie des Besitzes einer Internet-Flatrate (F, mit 0 = nicht vorhanden, 1 = vorhanden) dar. Der dazugehörige Datensatz umfasst 110 Werte. Wird zunächst nur der Einfluss des Entgelts (E) auf die Wahrscheinlichkeit der Teilnahme $p(Y)$ betrachtet, ergibt sich das in Abbildung C.28 angegebene Streuungsdiagramm. Die Werte wurden dabei in horizontaler Richtung grafisch gestreckt. Das Strecken ist lediglich als Hilfsmittel zu verstehen, so dass alle Datenpunkte sichtbar werden.

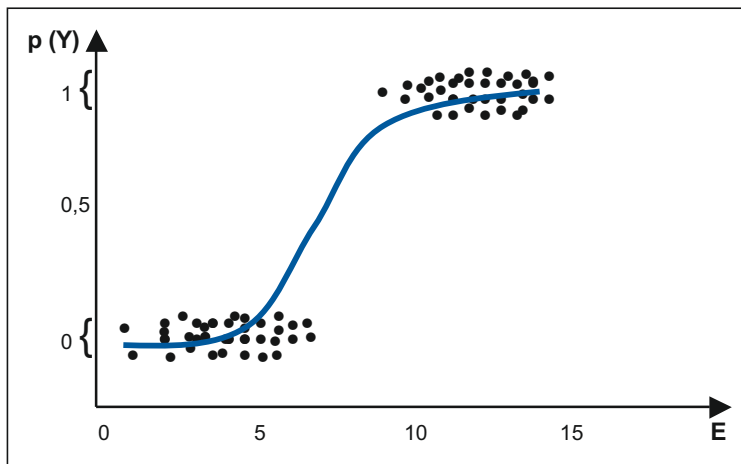


Abbildung C.28: Logistischer Funktionsverlauf

Die Grafik lässt vermuten, dass die Höhe des Entgelts (E) einen positiven Einfluss auf die Teilnahmewahrscheinlichkeit hat, da bei niedrigerer Entlohnung mehr Werte bei 0 liegen und umgekehrt. Um nun die in der Grafik dargestellte logistische Funktion zu schätzen, wird die Maximum-Likelihood-Schätzung (ML) eingesetzt. Sie wählt die Koeffizienten so, dass die Wahrscheinlichkeit der Beobachtung der empirischen Daten maximiert wird (Krafft, 2000, S. 242). Sie liefert für das multiple Modell das in Abbildung C.29 aufgeführte Ergebnis.

Variablen	Koeffizient	Exp (Koeffizient)	Standard- fehler	t-Wert	p-Wert
E	1,335	3,800	0,144	9,298	<0,001
F	0,743	2,102	0,290	2,562	0,012
Konstante	-9,175		1,031	-8,901	<0,001

Abbildung C.29: Schätzergebnisse für die logistische Regression

Die p-Werte des t-Tests (analog zur linearen Regression) weisen für alle Parameter signifikante Werte auf. Die Wahrscheinlichkeit der Teilnahme am Panel wird demnach laut (3) beschrieben durch:

$$\hat{p}(Y) = \frac{1}{(1 + \exp(-1 \cdot (-9,175 + 1,335 \cdot E + 0,743 \cdot F)))}$$

Mit den Schätzwerten lassen sich nun die Wahrscheinlichkeiten der Teilnahme am Panel berechnen. Für einen Probanden, der keine Internet-Flatrate besitzt ($F = 0$) und für die Teilnahme 5 Euro ($E = 5$) bekommt, gilt:

$$\hat{p}(Y | E = 5, F = 0) = \frac{1}{(1 + \exp(-1 \cdot (-9,175 + 1,335 \cdot 5 + 0,743 \cdot 0)))} = 0,076$$

Dieser Wert entspricht den odds von $0,076 / (1 - 0,076) = 0,082$.

Wird das Entgelt auf 6 Euro erhöht (c.p.), so ergibt sich analog eine vorhergesagte Wahrscheinlichkeit der Teilnahme von $\hat{p}(Y | E = 6, F = 0) = 0,238$ bzw. ein odds von $0,238 / (1 - 0,238) = 0,312$.

Wird der Quotient aus beiden odds gebildet, erhält man das so genannte **odds ratio** (OR).

Das odds ratio gibt den Faktor an, um den sich die Chance der Teilnahme ändert, sofern die unabhängige Variable um eine Einheit erhöht wird und alle weiteren Variablen konstant gehalten werden.

Für das Beispiel erhält man ein odds ratio von $OR = 0,312 / 0,082 = 3,8$. Eine Erhöhung des Entgelts für die Teilnahme um einen Euro steigert demnach die Chance, dass der Proband an der Umfrage teilnimmt um den Faktor 3,8. Dieser Wert entspricht $\exp(\text{Koeffizient})$ und wird wegen der besseren Interpretation häufig in den Schätzergebnissen mit aufgeführt. Für Probanden, die eine Internet-Flatrate besitzen, kann eine um den Faktor 2,1 höhere Chance an allen Wellen des Panels teilzunehmen, angenommen werden (Abbildung C.30).

Die **Güte** einer logistischen Regression kann zum einen anhand der Likelihood-Werte der Schätzung beurteilt werden. Ein Gütekriterium, das auf den Likelihood-Werten basiert, ist z. B. McFaddens R^2 (Krafft, 2000, S. 246). Es kann näherungsweise analog zum Bestimmtheitsmaß der linearen Regressionsanalyse interpretiert werden. Ein weiteres Kri-

		Prognostiziert		
		Teilnahme	Keine Teilnahme	Korrekte Klassifizierung
Beobachtet	Teilnahme	49	16	75,4 %
	Keine Teilnahme	14	31	68,9 %
Gesamtprozentsatz				72,7 %

Abbildung C.30: Klassifikationsmatrix

terium zur Beurteilung der Güte stellt der Anteil der durch das Modell richtig klassifizierten Fälle dar. Werden Wahrscheinlichkeitswerte von $p(Y) > 0,5$ als Teilnahme und Werte darunter als Nicht-Teilnahme interpretiert, so lassen sich diese Vorhersagen mit den tatsächlichen Werten vergleichen. Es ergibt sich die in Abbildung C.30 dargestellte Klassifikationsmatrix (allgemein Krafft, 2000, S. 246 f.). Der Anteil insgesamt korrekt klassifizierter Fälle, der auch als **hit ratio** bezeichnet wird (Krafft, 2000, S. 247), entspricht im Beispiel demnach $(49 + 31) / 110 = 0,727$, also 72,7 %.

Kausalanalyse

Wie die Regressionsanalyse ist auch die Kausalanalyse ein Verfahren zur empirischen Überprüfung von **Ursache-Wirkungs-Zusammenhängen** (Hildebrandt/Homburg, 1998). Im Vergleich zur Regressionsanalyse weist die Kausalanalyse jedoch insbesondere die folgenden Vorteile auf:

- Während die Regressionsanalyse eine fehlerfreie Messung der Variablen unterstellt, erlaubt die Kausalanalyse die explizite Berücksichtigung von Messfehlern.
- Die Kausalanalyse ermöglicht im Gegensatz zur Regressionsanalyse die Überprüfung komplexer Abhängigkeitsstrukturen, wie z. B. wechselseitige Abhängigkeiten (z. B. Markeneinstellung $\leftrightarrow$ Markenwahl) oder kausale Ketten (z. B. Werbung $\rightarrow$ Markenbekanntheit $\rightarrow$ Kaufabsicht).
- Korrelationen können zwischen den unabhängigen Variablen explizit im Rahmen der Modellformulierung und -schätzung berücksichtigt werden. Sie stellen somit im Gegensatz zur Regressionsanalyse (Multikollinearitätsproblem) kein prinzipielles Problem dar.

Die Kausalanalyse überwindet verschiedene Restriktionen der Regressionsanalyse, indem sie Messfehler explizit berücksichtigt, komplexe Abhängigkeitsstrukturen überprüft und Korrelationen zwischen den unabhängigen Variablen prinzipiell zulässt.

Zur Veranschaulichung der Kausalanalyse soll auf eine Studie zurückgegriffen werden, die sich mit den Erfolgsfaktoren von Markentransfers beschäftigt (Völckner, 2003; siehe hierzu genauer Kapitel E. 3.). Die Studie basiert auf einer Konsumentenbefragung (Quo-

tenstichprobe, $n = 2426$). Die Aufgabe der Befragten bestand darin, eine Reihe von Markentransfers anhand verschiedener Merkmale zu beurteilen. Dabei wurden unter anderem die Größen Muttermarkenstärke, Marketingunterstützung des Transferprodukts und subjektiv wahrgenommener Erfolg des Transferprodukts berücksichtigt. Die genannten Größen stellen **komplexe Konstrukte** dar, die nicht direkt beobachtet und gemessen werden können.

Die Kausalanalyse ist in der Lage, Wirkungszusammenhänge zwischen **komplexen Konstrukten** zu überprüfen. Eine grundlegende Besonderheit der Kausalanalyse liegt in der Unterscheidung von **beobachteten** (d. h. direkt messbaren) Variablen (so genannten Indikatorvariablen bzw. Items, z. B. die subjektiv wahrgenommene Qualität der Muttermarke) und **latenten Variablen** (z. B. das Konstrukt Muttermarkenstärke). Letztere stellen komplexe Größen dar, die nicht direkt gemessen werden können, sondern nur indirekt über die Indikatorvariablen erfasst werden können.

Die Grundidee der Kausalanalyse besteht nun darin, auf der Basis empirisch gemessener Varianzen und Kovarianzen von Indikatorvariablen Rückschlüsse auf die Wirkungszusammenhänge zwischen den zugrunde liegenden latenten Variablen zu ziehen (Homburg/Pflesser, 2000b, S. 643 f.). Die Analyse kann in vier Ablaufschritte unterteilt werden, die im Folgenden näher betrachtet werden sollen.

1. Schritt: Modellformulierung

Ein vollständiges Kausalmodell besteht aus jeweils einem Messmodell für die latenten exogenen (d. h. erklärenden) und latenten endogenen (d. h. durch die Kausalstruktur erklärten) Variablen und aus einem Strukturmodell.

Im Rahmen des **Messmodells** wird festgelegt, welche Konstrukte betrachtet und wie diese gemessen werden sollen. In der Regel wird eine einzelne latente Variable über mehrere Indikatorvariablen (Items) gemessen. Dabei wird unterstellt, dass jeder Indikator eine fehlerbehaftete Messung der zugrunde liegenden latenten Variablen darstellt (Hildebrandt/Homburg, 1998). Das Messmodell der Kausalanalyse entspricht damit dem Grundgedanken der konfirmatorischen Faktorenanalyse (siehe hierzu auch Kapitel C. 4.2). In Abbildung C.31 ist das Messmodell für das Konstrukt Muttermarkenstärke des Anwendungsbeispiels dargestellt. Auf einer 7-stufigen Ratingskala wurde jeweils der Grad der Zustimmung zu den einzelnen Items erfasst. Analog sind Messmodelle für die übrigen Konstrukte zu formulieren.

Ein Kausalmodell kann ‚reflektive‘ und ‚formative‘ Konstrukte enthalten. Das Kausalmodell in Abbildung C.31 umfasst ausschließlich **reflektive Konstrukte**. Ein reflektives Konstrukt ist dadurch gekennzeichnet, dass es die ihm zugeordneten Indikatoren verursacht. Die Beziehungen zwischen Konstrukt und zugehörigen Indikatoren werden hier über Faktorladungen beschrieben. Daneben gibt es **formative Konstrukte**, deren wesentliches Merkmal darin zu sehen ist, dass sie durch ihre Indikatoren verursacht werden. Die Indikatoren sind nicht austauschbar, sondern in ihrer Gesamtheit zur vollständigen Erfassung des betrachteten Konstrukts erforderlich (Chin, 1998).

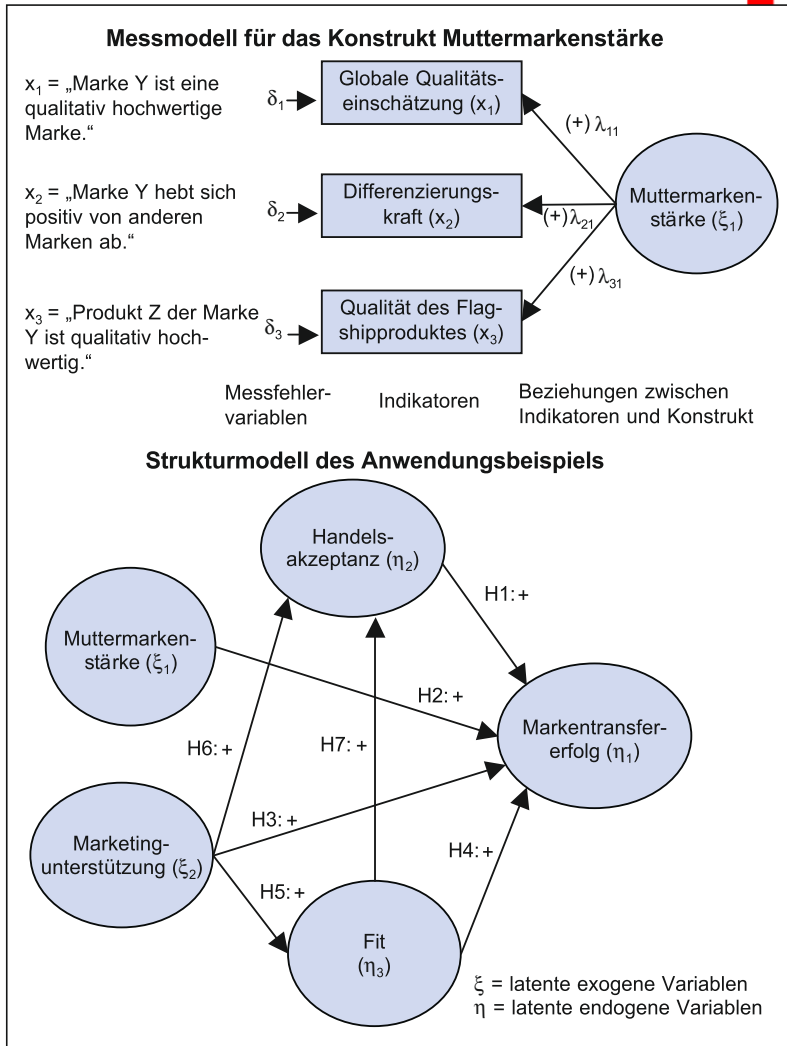


Abbildung C.31: Messmodell und Strukturmodell der Kausalanalyse

Das **Strukturmodell** beschreibt die hypothetischen Beziehungen zwischen den betrachteten Konstrukten (also den latenten Variablen). Die entsprechenden Hypothesen des Anwendungsbeispiels zwischen den Konstrukten sind in Abbildung C.31 dargestellt. Beispielsweise wurden hier die Hypothesen aufgestellt, dass die Marketingunterstützung einen positiven Einfluss auf die Handelsakzeptanz ausübt (Hypothese 6) und sich die wahrgenommene Ähnlichkeit zwischen Muttermarke und Transferprodukt positiv auf den Erfolg eines Transferprodukts auswirkt (Hypothese 4). Im Strukturmodell finden sich somit Grundgedanken der Regressionsanalyse wieder.

Reflektive und formative Konstrukte

Die Unterscheidung zwischen reflektiven und formativen Konstrukten soll anhand des Konstrukts „Trunkenheit“ veranschaulicht werden (Chin, 1998). Möchte man das Ausmaß der „Trunkenheit“ einer Person messen, so könnte man z. B. die Indikatoren Blutalkohol, Fahrtüchtigkeit und das Abschneiden der Person bei Kopfrechenaufgaben erfassen. „Trunkenheit“ stellt in diesem Fall ein reflektives Konstrukt dar. Denn je größer das Ausmaß der „Trunkenheit“ ist, desto schlechter wird z. B. die Fahrtüchtigkeit der Person oder ihr Abschneiden beim Kopfrechnen sein, d. h. das Konstrukt verursacht die zugrunde liegenden Indikatoren. Umgekehrt verursacht z. B. die konsumierte Menge an Bier, Wein und harten Spirituosen die „Trunkenheit“ einer Person. Das Konstrukt wird in letzterem Fall also durch die betrachteten Indikatoren verursacht und folglich als formatives Konstrukt gemessen.

Die Kausalanalyse stellt eine multivariate Methode dar, die Elemente der klassischen Regressionsanalyse und der konfirmatorischen Faktorenanalyse miteinander verbindet.

2. Schritt: Schätzung der Modellparameter

Zu den Modellparametern gehören:

- die Parameter, welche die Beziehungen zwischen den latenten Variablen beschreiben,
- die Koeffizienten der Pfade zwischen den latenten Variablen und ihren Indikatoren und
- die Kovarianzen der latenten exogenen Variablen, der Messfehlervariablen und der Residualvariablen der latenten endogenen Variablen (letztere sind mit der Residualgröße der Regressionsanalyse vergleichbar).

Die **Datengrundlage** für die Bestimmung der Modellparameter bilden die Varianzen und Kovarianzen der Indikatoren. Die in der Literatur weit verbreitete Bezeichnung Kausalanalyse ist insofern als problematisch anzusehen, da sie suggeriert, dass mithilfe eines statistischen Verfahrens Kausalität nachgewiesen werden könnte. Die Kovarianz stellt aber lediglich ein statistisches Kriterium dar, das eine Quantifizierung der zwischen den betrachteten Variablen bestehenden Beziehung erlaubt. Sie lässt keine Aussage darüber zu, welche der beiden Variablen als verursachend für die andere Variable anzusehen ist.

Bei der Kausalanalyse werden Korrelationen bzw. Kovarianzen auf verschiedene Weise kausal interpretiert. Die Aufdeckung von Ursache-Wirkungs-Zusammenhängen ist hingegen ausschließlich durch ein Experiment möglich (siehe hierzu Kapitel C. 4.5).

Das Ziel der Modellschätzung (typischerweise in Form einer Maximum-Likelihood-Schätzung) besteht darin, die Parameter so zu bestimmen, dass die auf Basis der Ergebnisse berechnete modelltheoretische Kovarianzmatrix $\Sigma = \Sigma(\alpha)$ der auf Basis der Stichprobe ermittelten Kovarianzmatrix S **möglichst ähnlich** wird. Dabei bezeichnet α den Vektor der zu schätzenden Parameter und $\Sigma(\alpha)$ die Kovarianzmatrix der Indikatoren als Funktion von α .

Im Zusammenhang mit der Parameterschätzung ist anzumerken, dass die Kausalanalyse ein Verfahren darstellt, das auf asymptotischer Statistik beruht. Daher ist eine sinnvolle Anwendung des Verfahrens ausschließlich bei hinreichend **großen Stichproben** möglich. Wie groß eine Stichprobe sein muss, lässt sich nicht pauschal beantworten, sondern hängt vor allem von der Modellkomplexität und der verwendeten Schätzmethode ab. In der Literatur wird für ML-Schätzungen häufig ein Verhältnis zwischen Stichprobenumfang und Anzahl zu schätzender Parameter von mindestens 5:1 empfohlen, sofern zumindest annähernd multivariat normalverteilte Variablen vorliegen (Bagozzi/Yi, 1988, S. 82).

In Abbildung C.32 sind die Ergebnisse der Parameterschätzung für die Beziehungen zwischen den latenten Variablen des Anwendungsbeispiels dargestellt. Die angegebenen Werte stellen analog zu den Beta-Koeffizienten der klassischen Regressionsanalyse standardisierte Schätzer dar, die unmittelbar miteinander verglichen werden können.

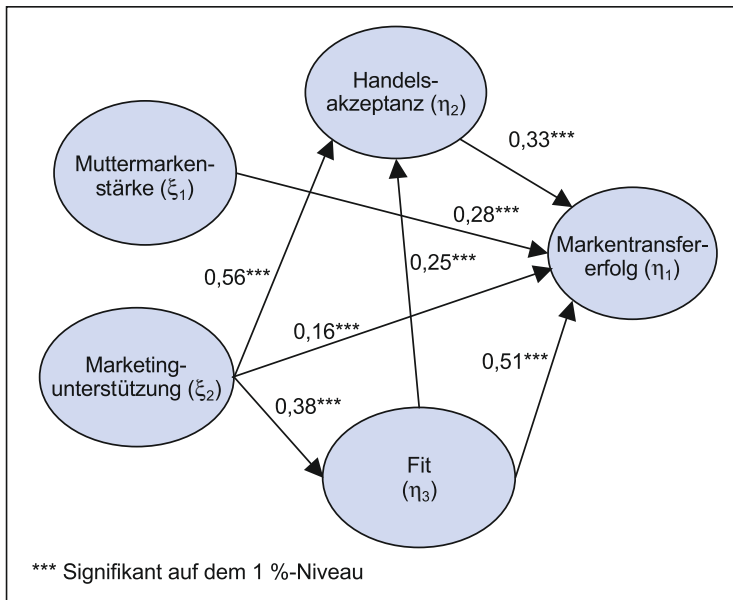


Abbildung C.32: Ausgewählte Ergebnisse der Parameterschätzung

Im Anschluss an die Schätzung der Modellparameter erfolgt eine Plausibilitätsbetrachtung der Ergebnisse. Dabei wird geprüft, ob nicht-sinnvolle Resultate auftreten, wie z. B. negative Varianzen oder Korrelationskoeffizienten größer als 1. Solche Werte liefern einen Hinweis darauf, dass ein Modell fehlerhaft spezifiziert wurde oder in Teilen nicht identifizierbar ist.

3. Schritt: Beurteilung der Modellgüte

Liegen keine unplausiblen Schätzwerte vor, so erfolgt im nächsten Schritt die Beurteilung der **Modellgüte** mithilfe von Anpassungsmaßen, welche die Güte der Anpassung des betrachteten Modells an den vorliegenden Datensatz beurteilen (z. B. Völckner, 2003).

Grundsätzlich kann zwischen **globalen** und **lokalen Anpassungsmaßen** unterschieden werden:

- **Globale Anpassungsmaße** geben Aufschluss darüber, wie gut die in den Hypothesen aufgestellten Beziehungen insgesamt durch die empirischen Daten wiedergegeben werden. Sie beruhen auf einem Vergleich der empirischen Kovarianzmatrix mit der modelltheoretischen Kovarianzmatrix.
- Mit **lokalen Anpassungsmaßen** werden die einzelnen Teilstrukturen des Modells beurteilt, d. h. die einzelnen Konstrukte und Indikatoren.

4. Schritt: Ergebnisinterpretation

Bei der Ergebnisinterpretation sollte sowohl eine Interpretation der Modellstruktur in ihrer Gesamtheit als auch eine Interpretation der einzelnen Koeffizienten hinsichtlich ihrer Signifikanz und Stärke erfolgen. Besonders aufschlussreich ist dabei die Betrachtung von direkten und indirekten Effekten der Konstrukte. Letztere treten bei der Analyse komplexer Abhängigkeitsstrukturen auf und können somit bei der klassischen Regressionsanalyse nicht erfasst werden

- **Direkte Effekte:** Im Anwendungsbeispiel geht der stärkste **direkte** Einfluss auf den Markentransfererfolg vom wahrgenommenen Fit zwischen Muttermarke und Transferprodukt aus. Der standardisierte Pfadkoeffizient in Höhe von 0,51 besagt, dass mit der Zunahme des wahrgenommenen Fit um eine Einheit die Erfolgswahrscheinlichkeit eines geplanten Markentransfers um 0,51 Einheiten zunimmt (der indirekte Effekt des Fit ist hier noch nicht berücksichtigt). Die Handelsakzeptanz übt mit einem Wert von 0,33 den zweitstärksten direkten Effekt auf den Markentransfererfolg aus.
- **Indirekte Effekte:** Neben den direkten Effekten liegen verschiedene **indirekte Einflüsse** der einzelnen Erfolgsfaktoren auf den Markentransfererfolg vor. So geht von der Marketingunterstützung nicht nur ein direkter Effekt auf den Markentransfererfolg aus, sondern es existieren auch indirekte Effekte über den Fit und die Handelsakzeptanz. Ein indirekter Effekt wird durch Multiplikation der einzelnen Effekte des indirekten Weges berechnet.
- **Totale Effekte:** Der **totale** Beeinflussungseffekt der Marketingunterstützung ergibt sich als Summe aus indirekten Effekten und direktem Effekt: $0,38 \cdot 0,25 \cdot 0,33 + 0,38 \cdot 0,51 + 0,56 \cdot 0,33 + 0,16 = 0,57$.

Dieses Beispiel veranschaulicht, dass durch die bloße Betrachtung der direkten Effekte die Einflussstärken der betrachteten Variablen unter Umständen erheblich unter- oder überschätzt werden. Die Kausalanalyse ist in der Lage, komplexe Wirkungszusammenhänge abzubilden und auf diese Weise auch die indirekten Effekte von Variablen offen zu legen.

4.4 Präferenzen analysieren

Präferenzen stellen einen eindimensionalen Indikator zur Erklärung von (Kauf-)Entscheidungen dar. Die Präferenz bringt das Ausmaß der Vorziehenswürdigkeit eines Beurteilungsobjekts für eine bestimmte Person während eines bestimmten Zeitraumes zum