

Auktionen versus Lotterien:

Ein empirischer Vergleich zur Messung von Zahlungsbereitschaften

Martin Schreier* und Joseph Werfer*

Working Paper

A later version of this paper is published in

Die Betriebswirtschaft (DBW)

*Abteilung für Entrepreneurship und Innovation

Wirtschaftsuniversität Wien

Augasse 2-6

A-1090 Wien

Email:

martin.schreier@wu-wien.ac.at

joseph.werfer@wu-wien.ac.at

Tel.: ++43/1/31336-5970

Fax: ++43/1/31336-769

Die Autoren danken Prof. Dr. Nikolaus Franke; Mag. Marion Pötz und Prof. Dr. Werner Müller für kritische Kommentare und Unterstützung, der Firma Apple für die Ermöglichung der realen Versteigerung der Produkte und der Wirtschaftsuniversität Wien für finanzielle Förderung des Projektes.

Auktionen versus Lotterien:

Ein empirischer Vergleich zur Messung von Zahlungsbereitschaften

Abstract (Deutsch)

Die verlässliche und valide Messung von Zahlungsbereitschaften ist für viele betriebswirtschaftliche Fragestellungen relevant (z.B. Pricing von neuen Produkten). Die beiden Methoden der Kaufangebote (a) Vickrey Auktionen und (b) Lotterien nach Becker, DeGroot und Marschak gelten in der Literatur als die besten. Ein empirischer Vergleich dieser beiden Methoden hinsichtlich der Qualitätskriterien Verständnis, strategisches Verhalten, Reliabilität und Validität fehlt bis dato. Der vorliegende Beitrag zielt darauf ab, diese relevante Lücke zu schließen.

Abstract (English)

A reliable and valid measurement of consumer's willingness to pay is relevant for many questions in the field of business administration (e.g. pricing for new products). The two methods (a) Vickrey auctions and (b) Becker, DeGroot and Marschak's lotteries are considered to be outstanding as opposed to their alternatives. An empirical comparison of these two methods with regard to issues of general understanding, reliability, and validity are missing to date. This article aims to fill this relevant gap.

1. Einführung

Ob ein Produkt am Markt überlebt, hängt mitunter von seinem Preis ab (Smith (1986)): Er trägt dazu bei, ob ein Konsument kauft oder nicht (Dodds/Monroe/Grewal (1991), Maxwell (2002), Simon (1992), S.4). Die Festlegung des Preises spielt demnach eine gewichtige Rolle bei der Einführung neuer Produkte. Ist ein Produkt im Regal, ist es schwierig und kostspielig, den Verkaufspreis im Nachhinein zu ändern (vgl. z.B. Simon (1992), S.656ff). Eine falsche Preisstrategie birgt problematische Folgerungen: Ist der Preis zu hoch, ist im Grenzfall kein Kunde bereit zu kaufen. Ist der Preis zu niedrig, bleibt ein Teil des Abschöpfungspotenzials ungenutzt. Aus Unternehmenssicht ist es folglich wichtig, a priori eine möglichst gute Einschätzung der durchschnittlichen Zahlungsbereitschaft (des Zielsegmentes) für ein neues Produkt zu bekommen (vgl. z.B. Monroe (2003), Noble/Gruca (1999)).

Die hohe praktische Relevanz der Messung von Zahlungsbereitschaften erklärt, warum auch die wissenschaftliche Forschung zu diesem Themenkomplex seit vielen Jahrzehnten aktiv arbeitet.¹ Zahlreiche Methoden wurden mit dem Ziel entwickelt, die Zahlungsbereitschaft von

¹ Vor allem Ökonomen und Psychologen, aber auch immer mehr Marktforscher, verwenden aggregierte Daten über Zahlungsbereitschaften um etwa die Nachfrage von Gütern zu schätzen (Wertenbroch/Skiera (2002), S.228). Anwendungsbeispiele reichen von Studien über Zahlungsbereitschaften zu umweltschonender Elektrizität („green energy“, siehe Roe et al. (2001), Rowlands et al. (2002)) bis zu Untersuchungen der Nachfrage im Gesundheitswesen (Blumenschein et al. (2001), Ryan/San Miguel (2000)). Skuras/Vakrou (2002) untersuchten etwa weiters den Einfluss der Herkunft eines Weines auf dessen Zahlungsbereitschaft, Hoffman et al. (1993) den Einfluss der Verpackung von Steaks, Loureiro et al. (2002) die Zahlungsbereitschaft für Öko-gekennzeichnete Äpfel und Franke/Piller (2004) bzw. Schreier (2005) den Wert, den selbst designte Produkte via „User-Toolkits“ für den Endkonsumenten generieren.

Kunden für ein ausgewähltes Produkt verlässlich und valide abzubilden. Bestehende Forschungsarbeiten haben gezeigt, dass hierzu besonders zwei Methoden der realen Kaufangebote geeignet sind (Auktionen nach Vickrey (1962) und Lotterien nach Becker/DeGroot/Marschak (1964)). Ein empirischer Vergleich dieser beiden Methoden fehlt bisher. Der vorliegende Beitrag zielt darauf ab, diese praktisch und gleichsam wissenschaftlich relevante Lücke zu schließen. Ist etwa eine der beiden Methoden das validere Instrumentarium zur Abbildung von Zahlungsbereitschaften, ist sie c.p. zu bevorzugen.

Im nächsten Abschnitt werden einerseits die gängigsten Möglichkeiten der Messung von Zahlungsbereitschaften und andererseits das Ziel der vorliegenden Arbeit skizziert. Im Anschluss wird die Vorgehensweise für die empirische Erhebung genau dargelegt. Darauf folgt die Präsentation der Ergebnisse – insbesondere hinsichtlich der Reliabilität bzw. Validität der beiden Methoden zur Erfassung von Zahlungsbereitschaften. Der Beitrag schließt mit einer Zusammenfassung und einem Ausblick für zukünftige Forschungsarbeiten.

2. Hintergrund und Forschungslücke

2.1 Möglichkeiten der Messung von Zahlungsbereitschaften

Zahlungsbereitschaften beziffern jeweils den maximalen Geldbetrag, den potenzielle Kunden für ein ausgewähltes Produkt zu bezahlen bereit sind (Sattler/Nitschke (2003), S. 364, Simon (1992), S.90, Skiera/Revenstorff (1999), S.224). Sie können grundsätzlich in Form von (1) Kaufdaten, (2) Präferenzdaten oder (3) Kaufangeboten erhoben werden (Sattler/Nitschke (2003), S.365).

(1) Kaufdaten werden im realen Kaufumfeld erfasst. Die Durchführung einer derartigen Erhebung erfolgt entweder im Rahmen eines experimentellen (z.B. Testmarktsetting) oder eines nicht-experimentellen Settings (z.B. Scanner-Paneldaten). Kerninformation ist hierbei, ob Kunden tatsächlich bereit sind, zu einem ausgewiesenen Preis zu kaufen oder nicht (dichotome Information zur Zahlungsbereitschaft pro Erhebung bzw. Experiment). Die wahre, individuelle Zahlungsbereitschaft bleibt bei dieser Methode verborgen. Der hohen externen Validität steht daher eine niedrige Varianz der Zahlungsbereitschaften gegenüber (Ben-Akiva et al. (1994), S.344, Sattler/Nitschke (2003), S.364).

(2) Präferenzdaten können mit Hilfe von direkten bzw. indirekten Preisbefragungen erhoben werden (Sattler/Nitschke (2003), S.364). Erstere richtet z.B. die direkte Frage an Probanden: „Wieviel würden Sie maximal für das Produkt X bezahlen“ (diese und alternative Formen der direkten, hypothetischen Preisbefragung sind als „Contingent Valuation Methods“ bekannt – CVM, vgl. Carson et al. (1995), Mitchell/Carson (1989)). Voraussetzung für eine valide Messung ist, dass Befragte einen Anreiz haben, ihre wahre Zahlungsbereitschaft bekanntzugeben. Dies ist allerdings nicht immer der Fall (Fox et al. (1998), List/Shogren (1998), Neill et al. (1994)). List/Gallet (2001) haben etwa gezeigt, dass hypothetische Methoden die wahre Zahlungsbereitschaft um den Faktor >3 überschätzen.

Die bekannteste Form der indirekten Preisbefragung ist die Conjoint-Analyse (Green/Srinivasan (1990) - bezüglich neuerer Ansätze siehe z.B. Dahan/Hauser (2002)). „Die Aufgabe der Befragten besteht üblicherweise darin, ein Rating, Ranking oder eine Wahlentscheidung zwischen alternativen Produktkonzepten vorzunehmen, die sich systematisch hinsichtlich Preis und anderer Produkteigenschaften unterscheiden.“ (Sattler/Nitschke (2003), S.365). Die Conjoint-Analyse basiert auf der Annahme, dass man

Präferenz- in Kaufdaten transformieren kann. Auch diese Annahme ist nicht immer unproblematisch (Ben-Akiva et al. (1994), S.344). Darüber hinaus bleibt die Conjoint-Analyse „wirkungslos“, wenn keine Produkt- bzw. Produktattributalternativen zur Verfügung stehen (wenn also die Zahlungsbereitschaft eines einzigen, gesamten Produktes ermittelt werden soll).

(3) Das Offerieren von Kaufangeboten bildet die dritte Form zur Abbildung von Zahlungsbereitschaften. In Form von (a) Auktionen oder (b) Lotterien bekommen Probanden die Möglichkeit, die zugrunde liegenden Produkte tatsächlich zu erwerben (*reale* Messung von Zahlungsbereitschaften im Vergleich zur *hypothetischen* Messung im Rahmen der CVM). Hinsichtlich ersterer Variante hat sich der Auktionstyp nach William Vickrey (Vickrey (1962)) als geeignetste Form etabliert (bezüglich alternativer Auktionsformen siehe z.B. Wolfstetter (1996)). In verdeckter Form gibt jeder Teilnehmer sein Gebot schriftlich ab. Sieger der Auktion ist der Bieter mit dem höchsten Gebot. Im Gegensatz zur verdeckten Höchstpreisauktion muss dieser aber nicht die Höhe seines Gebotes bezahlen. Der Kaufpreis richtet sich stattdessen an das zweithöchste Gebot (bzw. an das n -höchste Gebot). Dies bedeutet, dass die Bieter mit ihrem Gebot den Preis nicht direkt beeinflussen können. Die dominante Strategie der Vickrey Auktion ist demnach ein Gebot in Höhe der wahren Zahlungsbereitschaft abzugeben (Vickrey (1962)). Bietet man mehr als seine maximale Zahlungsbereitschaft, läuft man Gefahr die Auktion zu einem Preis zu gewinnen (in der Höhe des zweit- bzw. n -höchsten Gebots), der über der maximalen Zahlungsbereitschaft liegt (und man erleidet dadurch einen Verlust). Bietet man weniger als die maximale Zahlungsbereitschaft, besteht das Risiko, dass die Auktion von einem anderen Bieter gewonnen wird, obwohl der Preis (in der Höhe des zweit- bzw. n -höchsten Gebots) niedriger ist als die eigene maximale Zahlungsbereitschaft (und man erleidet ebenfalls einen Verlust) (Wolfstetter (1996), S.371ff, Skiera/Revenstorff (1999)). Diese strategische

Rahmenbedingung wird als Anreizkompatibilität bezeichnet (Kagel (1995)). Dies unterscheidet die Vickrey Auktion auf theoretischer Ebene sowohl von anderen Auktionsformen als auch von der direkten, hypothetischen Preisbefragung.

Im Zuge von Lotterien nach Becker/DeGroot/Marschak (1964) (im folgenden BDM) geben Probanden zunächst ihre Zahlungsbereitschaft für ein konkretes Produkt bekannt. Im Anschluss wird zufällig ein Preis gezogen. Liegt die Zahlungsbereitschaft über dem zufällig gezogenen Preis, *muss* der Proband kaufen, liegt sie unterhalb, *darf* er das Produkt *nicht* kaufen (Becker/DeGroot/Marschak (1964)). Wie die Vickrey Auktion, erfüllt auch diese Methode das Kriterium der Anreizkompatibilität. Die dominante Strategie des Lotterieteilnehmers ist es, die wahre Zahlungsbereitschaft bekanntzugeben. Darüber hinaus wird die BDM dem sogenannten „Point-of-Purchase“ Kriterium gerecht. Dies bedeutet, dass sich Teilnehmer während der Durchführung im gewohnten Verkaufsumfeld befinden können (Wertenbroch/Skiera (2002)). Dieses Argument spricht für die externe Validität der Methode.

2.2 Forschungslücke: ein empirischer Vergleich zwischen Vickrey Auktion und BDM

Sowohl Kauf- als auch Präferenzdaten werden von Unternehmen häufig zur Nachfrageeinschätzung und für preispolitische Entscheidungen eingesetzt. Praktische Anwendungsfälle der Methoden der Kaufangebote, z.B. in den Bereichen Marketing bzw. Innovationsmanagement, sind jedoch sehr bescheiden (Skiera/Revenstorff (1999)) – obwohl diesen Methoden einerseits theoretische und andererseits empirisch belegte Überlegenheit attestiert wurde. Skiera/Revenstorff (1999) haben etwa empirisch gezeigt, dass Vickrey Auktionen ein valides Instrument zur Abbildung von Zahlungsbereitschaften sind. Wertebroch und Skiera (2002) konnten dies für die BDM zeigen. Ein *direkter* empirischer Vergleich zwischen Vickrey Auktion und BDM fehlt jedoch bisher.

Der vorliegende Beitrag versucht diese Lücke zu schließen. Im Detail werden die beiden folgenden Fragestellungen bearbeitet:

(1) Gibt es signifikante Unterschiede zwischen Vickrey Auktion und BDM hinsichtlich der durchschnittlich erfassten Zahlungsbereitschaft? Beide Methoden sind theoretisch anreizkompatibel (dominante Strategie: Bekanntgabe der wahren Zahlungsbereitschaft). Empirisch dürften demnach keine signifikanten Unterschiede in der Zahlungsbereitschaft zwischen den Methoden zu beobachten sein.

(2) Welche Methode (Vickrey Auktion versus BDM) eignet sich besser, die Nachfrage reliabel und valide abzubilden?

Antworten auf diese Fragestellungen sind nicht nur aus wissenschaftlicher Perspektive relevant, sondern auch z.B. für Marktforscher und Innovationsmanager erfolgsweisend. Nur wenn Zahlungsbereitschaften reliable und valide erfasst werden können, ist der Einsatz z.B. für die Preisbildung von Produkten empfehlenswert. Stellt sich eine Methode als reliabler bzw. valider dar, ist sie c.p. zu bevorzugen. Gleiches gilt für wissenschaftliche Anwendungen. Franke/Schreier (2005) versuchten etwa den Wertzuwachs selbst designer Produkte im Vergleich zu herkömmlichen Standardprodukten mit Hilfe von Maximum-Likelihood Schätzungen zu erklären. Als abhängige Variable dienten Zahlungsbereitschaften - gemessen mit Hilfe von Vickrey Auktionen.

3. Beschreibung der Untersuchung

3.1 Untersuchungssetting

Probanden, die Interesse an einer Teilnahme der Untersuchung hatten (Akzeptanzquote von ca. 60%), wurde zunächst der Ablauf der Untersuchung erklärt. Im Anschluss konnten sie Informationsbroschüren studieren und das zu bewertende Produkt vor Ort testen. Darüber hinaus wurde ihnen das Produkt von fachkundigen Personen nähergebracht. Nach erfolgreichem Produkttest und nachdem Probanden genügend Informationen hatten, um den Wert des Produktes für sich einschätzen zu können, wurde die genaue Funktionsweise der Preisermittlung erklärt.

Die Zuteilung, welche Methode zur Erfassung der Zahlungsbereitschaft zur Anwendung kam, erfolgte nach einem Zufallsalgorithmus. Zusätzlich zu den beiden Methoden der Kaufangebote (Vickrey Auktion bzw. BDM) kam die direkte hypothetische Preisbefragung (CVM) als Kalibrierungsmethode zur Anwendung. Dadurch kann nicht nur der direkte Vergleich zwischen Vickrey Auktion und BDM sondern auch deren relative Reliabilitäts- bzw. Validitätsstärke im Vergleich zu einer in der Literatur schwächer eingestuften Methode angestellt werden (vgl. z.B. Wertenbroch/Skiera (2002)).

Jeder Proband reichte einen schriftlichen Gebotszettel ein (schriftliche Abgabe der Zahlungsbereitschaft mit bindendem Charakter; auf dem Gebotszettel befand sich außerdem eine schriftliche Erklärung der Funktionsweise der jeweiligen Methode). Potenzielle Interaktionen zwischen Probanden (z.B. Preisabsprachen) wurden bewusst und erfolgreich unterbunden. Abschließend füllten Probanden einen schriftlichen Fragebogen aus, der Reliabilitäts- bzw. Validitätsitems, Verständnisfragen und allgemeine Fragen zur Person beinhaltete (die Operationalisierung zur Reliabilitäts-, Validitäts- und Verständnisbeurteilung

erfolgte in Anlehnung an Sattler/Nitschke (2003), Skiera/Revenstorff (1999) und Wertenbroch/Skiera (2002), vgl. Tabelle 1).

Im Anschluss an die Erhebung kamen sowohl die Sieger im Rahmen der Vickrey Auktion als auch jene, die im Rahmen der BDM den Zuschlag erhielten, ihrer Kaufverpflichtung nach. Dies ist ein erstes Indiz, dass die reale Kaufverpflichtung verstanden und ernst genommen wurde.

Prüfkriterium	Operationalisierung / Messung	Skala
Verständnis bzw. strategisches Verhalten	Item 1: Ich will jetzt eigentlich keinen iPod. Deswegen ist mein Gebot so niedrig, dass ich auf keinen Fall den Zuschlag erhalte.	5-stufige Ratingskala (1=stimme gar nicht zu; 5=stimme völlig zu)
	Item 2: Ich habe auf einen besonders günstigen Kauf spekuliert.	5-stufige Ratingskala (1=stimme gar nicht zu; 5=stimme völlig zu)
	Item 3: Die Erklärungen zur Vorgehensweise waren für mich verständlich.	5-stufige Ratingskala (1=stimme gar nicht zu; 5=stimme völlig zu)
	Item 4: Ich bin mir bewusst, dass mein Gebot bindend ist.	5-stufige Ratingskala (1=stimme gar nicht zu; 5=stimme völlig zu)
	Item 5: Ich bin mir bewusst, dass es sich hier um eine reale Kaufsituation handelte. (Wertenbroch/Skiera (2002))	5-stufige Ratingskala (1=stimme gar nicht zu; 5=stimme völlig zu)
	Item 6a (Vickrey Auktion): Bieter A bietet €200, Bieter B bietet €300, und Bieter C €250. Wer erhält den Zuschlag? Wieviel muss bezahlt werden? (Sattler/Nitschke (2003))	Richtige Antwort: Bieter B zu €250.
	Item 6b (BDM): Bieter A bietet €200, Bieter B bietet €300, und Bieter C €250. Gezogen wird ein Preis von €230. Wer erhält den Zuschlag? Wieviel muss bezahlt werden?	Richtige Antwort: Bieter B und Bieter C zu €230.
Reliabilität	Vergleich der drei Tagessubsamples (z.B. Green/Tull/Albaum (1988), S.253, Wertenbroch/Skiera (2002))	
	Item 7: Hätte ich mir länger über die Höhe des Gebotes Gedanken gemacht, hätte ich mich vielleicht anders entschieden	5-stufige Ratingskala (1=stimme gar nicht zu; 5=stimme völlig zu)
Face Validität	Item 8: Hätte ich noch einmal die Chance ein Gebot abzugeben, würde ich es exakt in der gleichen Höhe abgeben.	5-stufige Ratingskala (1=stimme gar nicht zu; 5=stimme völlig zu)
	Korrelationsanalysen zwischen WTP und Produktinteresse und Kaufwahrscheinlichkeit (z.B. Wertenbroch/Skiera (2002), Skiera/Revenstorff (1999))	
	Item 9: Wie hoch ist Ihr Interesse an einem iPod?	5-stufige Ratingskala (1=gar nicht hoch; 5=sehr hoch)
	Item 10: Wie hoch ist die Wahrscheinlichkeit, dass Sie in den nächsten 6 Monaten einen iPod	5-stufige Ratingskala (1=gar nicht wahrscheinlich; 5=sehr wahrscheinlich)

kaufen werden?

Interne Validität	Vergleich beobachtete vs. geschätzte Nachfragefunktion (Logit-Regressionen) Vergleich von R^2 und Elastizität (Wertenbroch/Skiera (2002))	-
Externe Validität	Umlegbarkeit der WTP auf andere Situationen: Item 11: Wenn ich das Produkt jetzt sofort bekäme, würde ich mehr/weniger bezahlen. Item 12: In einem Onlineshop würde ich mehr/weniger bezahlen. Item 13: In einem herkömmlichen Geschäft würde ich mehr/weniger bezahlen.	6-stufige Ratingskala (-3=sehr viel weniger; 0=gleich viel; 3=sehr viel mehr) 6-stufige Ratingskala (-3=sehr viel weniger; 0=gleich viel; 3=sehr viel mehr) 6-stufige Ratingskala (-3=sehr viel weniger; 0=gleich viel; 3=sehr viel mehr)

Tabelle 1: Operationalisierung zur Reliabilitäts-, Validitäts- und Verständnisbeurteilung

3.2 Forschungsobjekt

Als Forschungsobjekt diente der *iPod* (MP3 Player) der Firma *Apple*. Dieser MP3 Player verfügt über eine 15GB-Festplatte und fasst bis zu 10.000 Lieder. Das Gerät hat in etwa die Größe einer Audiokassette und das Gewicht von zwei CDs. Der Verkaufspreis dieser Produktkategorie bewegt sich zwischen €200 und €400 (für nähere Details siehe www.apple.com).

3.3 Stichprobe

Grundgesamtheit der Untersuchung waren Studierende der Universität der Autoren. Probanden wurden direkt am Campus angesprochen. Insgesamt wurden 323 BWL-Studierende im Laufe von drei Tagen befragt. Für 109 Probanden kam die Vickrey Auktion, für 110 die BDM und für 104 die CVM zur Anwendung (randomisierte Zuteilung). Tabelle 2 gibt einen Überblick über die Charakteristika der Stichprobe(n). Der Anteil männlicher Probanden liegt bei 67%. Das Alter beträgt durchschnittlich 23,8 Jahre. Das verfügbare Monatseinkommen liegt bei 25% der Untersuchungsteilnehmer unter €200, für 36% zwischen

€201 und €400 und für 39% über €400. Unterschiede zwischen den drei Subsamples (Vickrey Auktion, BDM bzw. CVM) erweisen sich durchgängig als nicht signifikant. Auch hinsichtlich Produktinteresse und Informationsstand über den iPod, bzw. über MP3 Player generell, gibt es keine signifikanten Unterschiede zwischen den Subsamples (vgl. Tabelle 2). Unterschiede hinsichtlich der Höhe der Zahlungsbereitschaften zwischen den drei Subsamples sollten demnach auf die Methoden (und nicht auf unterschiedliche Stichprobensamensetzungen) zurückzuführen sein.

Variable	Vickrey Auktion (n=109)	BDM (n=110)	CVM (n=104)	Gesamt (n=323)
Geschlecht (Prozent m/w) ^a	62:38	70:30	61:39	65:35
Alter (Mittelwert) ^b	24,08	23,60	23,97	23,88
Einkommen (Prozent 3 Gruppen) ^c	29:34:37	21:38:41	21:44:35	27:37:36
Produktinteresse (MW; SA) ^d	2,95 (1,25)	2,98 (1,20)	3,08 (1,20)	3,00 (1,22)
Informationsstand Ipod (MW; SA) ^e	2,73 (1,48)	2,76 (1,46)	2,41 (1,42)	2,64 (1,46)
Informationsstand MP3 Player (MW; SA) ^f	3,72 (1,06)	3,74 (1,01)	3,55 (1,14)	3,67 (1,07)

*p<0,05 **p<0,01 ***p<0,001

^a χ^2 -Test: $\chi^2=2,677$; p=0,262

^b t-Tests: VA/BDM (t=1,101; p=0,272); VA/CVM(t=0,251; p=0,802); BDM/CVM(t=-0,779; p=0,437)

^c Verfügbares Einkommen pro Monat (in Euro); Gruppen: < 200; 201-400; > 400

t-Tests: VA/BDM (t=-0,789; p=0,431); VA/CVM(t=0,291; p=0,841); BDM/CVM(t=0,993; p=0,322)

^d Gemessen auf einer Skala von 1 bis 5 (1=sehr niedrig; 5=sehr hoch)

t-Tests: VA/BDM (t=-0,167; p=0,868); VA/CVM(t=-0,730; p=0,466); BDM/CVM(t=-0,578; p=0,564)

^e Gemessen auf einer Skala von 1 bis 5 (1=sehr niedrig; 5=sehr hoch)

t-Tests: VA/BDM (t=-0,150; p=0,881); VA/CVM(t=1,614; p=0,108); BDM/CVM (t=1,779; p=0,077)

^f Gemessen auf einer Skala von 1 bis 5 (1=sehr niedrig; 5=sehr hoch)

t-Tests: VA/BDM (t=-0,083; p=0,934); VA/CVM(t=1,136; p=0,257); BDM/CVM(t=1,244; p=0,215)

Tabelle 2: Charakteristika der Stichprobe

4. Befunde

4.1 Die ermittelten Zahlungsbereitschaften

Insgesamt gaben 323 Probanden ihre Gebote für den iPod ab. Die Zahlungsbereitschaften schwanken bei der Vickrey Auktion zwischen €0,50 und €350, bei der BDM zwischen €10 und €260 und bei der CVM zwischen €5 und €500 (vgl. Tabelle 3). Der Mittelwert beträgt bei der Vickrey Auktion €85 (Standardabweichung SA: 69), bei der BDM €73 (SA: 59) und bei der CVM ohne reale Kaufverpflichtung €164 (SA: 105).² Zwischen den beiden Methoden der Kaufangebote gibt es *keine* signifikanten Unterschiede ($p=0,176$). Bezüglich der Messung der durchschnittlichen Zahlungsbereitschaft sollte es (bei entsprechend großen Stichproben) demnach keine Rolle spielen, ob Vickrey Auktionen oder BDM Lotterien eingesetzt werden. Beide Methoden liefern aber einheitlich signifikant *niedrigere* Zahlungsbereitschaften als nach der CVM-Methode (p jeweils <0.001). Die hypothetische überschätzt somit im Vergleich zu den realen Methoden die Zahlungsbereitschaft etwa um das Doppelte.

Zahlungsbereitschaften	Vickrey Auktion ($n=109$)	BDM ($n=110$)	CVM ($n=104$)
Mittelwert (SA) ^a	84,68 (68,62)	72,96 (58,68)	164,43 (105,47)
Minimum / Maximum	0,50/350	10/260	5/500
Zahlungsbereitschaften (sortiert)	Kumulierte Häufigkeit/ Absolute Häufigkeit	Kumulierte Häufigkeit/ Absolute Häufigkeit	Kumulierte Häufigkeit/ Absolute Häufigkeit
0,5	0,9% 1	- -	- -
1	3,7% 3	- -	- -
5	8,2% 5	- -	1,0% 1
10	14,7% 7	2,7% 3	- -
15	19,2% 5	5,5% 3	1,9% 1
19,9	20,2% 1	- -	- -
20	22,9% 3	16,4% 12	3,8% 2
21	- -	17,3% 1	- -
25	23,8% 1	25,5% 9	4,8% 1
28	- -	26,4% 1	- -
30	25,7% 2	31,8% 6	7,7% 3
35	28,4% 3	36,4% 5	8,7% 1
40	31,2% 3	43,6% 8	9,6% 1
45	33,9% 3	- -	- -
49	- -	44,5% 1	- -
50	43,1% 10	53,6% 10	15,4% 6
59	44,0% 1	- -	- -

² Die deutlichen Abweichungen der durchschnittlichen Zahlungsbereitschaften aller drei Methoden vom Verkaufspreis im Handel ist nicht überraschend, da die Stichprobe nicht mit der Zielgruppe des iPods gleichgesetzt werden darf (vgl. Sattler/Nitschke (2003)). Wie in Tabelle 2 ersichtlich, ist etwa das Produktinteresse der Befragten im Schnitt „nur“ mittelmäßig (Mittelwert um den Wert 3 auf einer 5-stufigen Skala wobei 5 jeweils „stimme völlig zu“ entspricht) und weist außerdem eine relative hohe Varianz auf (Standardabweichung von durchschnittliche 1,2). Vgl. diesbezüglich auch Abschnitt 4.4.

60	48,6%	5	55,5%	2	17,3%	2
65	49,5%	1	56,4%	1	-	-
69	50,4%	1	-	-	-	-
70	55,0%	5	63,6%	8	21,2%	4
71	-	-	64,5%	1	-	-
75	-	-	65,5%	1	22,1%	1
80	56,9%	2	66,4%	1	23,1%	1
81	-	-	67,3%	1	-	-
90	-	-	70,0%	3	25,0%	2
95	57,8%	1	-	-	-	-
99	-	-	-	-	27,9%	3
100	67,0%	10	78,2%	9	35,6%	8
107	-	-	-	-	36,5%	1
110	68,8%	2	79,1%	1	-	-
111	68,8%	-	80,0%	1	-	-
120	72,5%	4	81,8%	2	38,5%	2
125	-	-	83,6%	2	39,4%	1
140	75,2%	3	-	-	41,3%	2
144	-	-	-	-	42,3%	1
145	76,1%	1	-	-	-	-
149	77,0%	1	-	-	-	-
150	90,8%	15	91,8%	9	58,7%	17
165	-	-	-	-	59,6%	1
170	92,6%	2	-	-	61,5%	2
180	-	-	92,7%	1	64,4%	3
190	93,6%	1	93,6%	1	-	-
199	-	-	-	-	65,4%	1
200	96,3%	3	97,3%	4	77,9%	13
221	96,3%	-	-	-	78,8%	1
240	97,2%	1	-	-	-	-
250	98,1%	1	99,1%	2	84,6%	6
260	-	-	100,0%	1	-	-
289	-	-	-	-	85,6%	1
300	99,1%	1	-	-	92,3%	7
350	100,0%	1	-	-	93,3%	1
370	-	-	-	-	94,2%	1
399	-	-	-	-	95,2%	1
400	-	-	-	-	97,1%	2
450	-	-	-	-	99,0%	2
500	-	-	-	-	100,0%	1

^a t-Tests: VA/BDM (t=1.359; p=0,176), VA/CVM (t=-6,571; **p=0,000**), BDM/CVM (t=-7,895; **p=0,000**)

Tabelle3: Verteilung der Zahlungsbereitschaften

4.2 Verständnis der Methoden und strategisches Verhalten

Zunächst wird untersucht, ob sich hinsichtlich des Verständnisses der Funktionsweisen der beiden Methoden Vickrey Auktion bzw. BDM Unterschiede identifizieren lassen (für die CVM sind die Verständnisfragen nicht anwendbar, vgl. Tabelle 4).

Im Zuge der Befragung wurde den Probanden mündlich und schriftlich die Vorgehensweise erklärt, bei den Methoden mit Anreizkompatibilität darüber hinaus ein schriftliches Beispiel gegeben. Insgesamt zeigen die Ergebnisse, dass die Vorgehensweise für die Probanden

größtenteils klar und verständlich war (Item3). Unterschiede zwischen Vickrey Auktion und BDM erweisen sich als nicht signifikant. Item 2 und 3 zeigen außerdem, dass den Probanden klar war, dass die Gebote verbindlich waren und dass es sich damit um eine reale Kaufsituation handelte. Der genaue Mechanismus der Methoden wurde jedoch nicht durchgängig verstanden. Für beide Methoden wurde eine Kontrollfrage gestellt (vgl. Tabelle 4). Bei der Vickrey Auktion haben ca. 15% die Frage und damit die Anreizkompatibilität nicht verstanden, bei der BDM waren es sogar ca. 23% - trotz des hohen Aufwandes, genau dies zu verhindern.³

Zusätzlich wurde mit zwei Items das strategische Verhalten abgefragt. Sowohl das Statement „Ich will jetzt eigentlich keinen iPod. Deswegen ist mein Gebot so niedrig, dass ich auf keinen Fall den Zuschlag erhalte“ als auch das Item „Ich habe auf einen besonders günstigen Kauf spekuliert“ erhalten für beide Methoden eine verhältnismäßig hohe Zustimmung (Mittelwerte von 3,20 bis 3,87 jeweils auf einer 5-stufigen Skala). Während ersteres Item höhere Zustimmung bei der Vickrey Auktion findet ($p > 0,05$), ist das letztere Item signifikant stärker bei der BDM ausgeprägt.

Korrelationsanalysen mit der Gebotshöhe zeigen, dass dieses strategische Verhalten die Höhe der Zahlungsbereitschaften beeinflusst. Die Angst vor einem möglichen Zuschlag korreliert sowohl mit den Geboten der Vickrey Auktion als auch mit denen der BDM ($r = -0,36^{**}$ bzw. $-0,21^{*}$; Unterschiede zwischen den Koeffizienten nach t-Test der Fisher z-Scores⁴ erweisen sich als nicht signifikant). Auch die Spekulation auf einen besonders günstigen Kauf stellt

³ 12,2% bei Sattler/Nitschke (2003)

⁴ Vgl. z.B. Glass/ Stanley (1970), Wertebroch/Skiera (2002)

einen Störfaktor dar (Vickrey Auktion: $r=-0,26^{**}$; BDM: $r=-0,14$ n.s.; Unterschiede zwischen den Koeffizienten sind signifikant; $p<0,01$).

Theoretisch (gemäß Anreizkompatibilität) sollten Probanden durchgängig genau maximal soviel bieten, wie ihnen das zugrunde liegende Produkt zum Zeitpunkt der Erhebung wert ist. Anders ausgedrückt: etwaige „Störeinflüsse“ wie strategisches Verhalten sollten keine Rolle spielen. In Summe weisen die Befunde dieser Untersuchung jedoch darauf hin, dass die theoretische Anreizkompatibilität empirisch nicht zwingend ihre vollständige Wirkung entfaltet (vgl. Sattler/Nitschke (2003)). Sowohl Verständnis als auch strategisches Bietverhalten stellen empirische Störfaktoren dar.

Strategische Überlegungen	Vickrey Auktion MW (SA) (<i>n=109</i>)	BDM MW (SA) (<i>n=110</i>)
Item 1: Ich will jetzt eigentlich keinen iPod. Deswegen ist mein Gebot so niedrig, dass ich auf keinen Fall den Zuschlag erhalte ^a	3,33(1,48)	3,20(1,41)
Item 2: Ich habe auf einen besonders günstigen Kauf spekuliert ^b	3,26(1,45)	3,87(1,23)
Item 3: Die Erklärungen zur Vorgehensweise waren für mich verständlich ^c	4,40(0,87)	4,46(0,80)
Item 4: Ich bin mir bewusst, dass mein Gebot bindend ist ^d	4,50(1,02)	4,64(0,90)
Item 5: Ich bin mir bewusst, dass es sich hier um eine reale Kaufsituation handelt ^e	4,52(0,95)	4,54(1,00)
Kontrollfrage Vickrey A Item 6a: Bieter A bietet €200, Bieter B bietet €300, und Bieter C €250. Wer erhält den Zuschlag? Wieviel muss bezahlt werden?	85,19% richtig	-
Kontrollfrage BDM Item 6b: Bieter A bietet €200, Bieter B bietet €300, und Bieter C €250. Gezogen wird ein Preis von €230. Wer erhält den Zuschlag? Wieviel muss bezahlt werden?	-	77,27% richtig

^a Gemessen auf einer Skala von 1 bis 5 (1=Stimme gar nicht zu; 5=Stimme völlig zu);
t-Tests: VA/BDM (t=0,656; p=0,513); Korrelation mit WTP: r=-0,361** (Vickrey Auktion); r=-0,214* (BDM);
t-Tests der Fisher z-Scores: t=-1,445; p=0,150

^b Gemessen auf einer Skala von 1 bis 5 (1=Stimme gar nicht zu; 5=Stimme völlig zu); Korrelation mit WTP:
r=-0,261** (Vickrey Auktion); r=-0,136 (BDM); t-Tests der Fisher z-Scores: t=3,094; p=0,002
t-Tests: VA/BDM (t=-3,386; **p=0,001**);

^c Gemessen auf einer Skala von 1 bis 5 (1=sehr niedrig; 5=sehr hoch);
t-Tests: VA/BDM (t=-0,531; p=0,596);

^d Gemessen auf einer Skala von 1 bis 5 (1=sehr niedrig; 5=sehr hoch);
t-Tests: VA/BDM (t=-1,050; p=0,295);

^e Gemessen auf einer Skala von 1 bis 5 (1=sehr niedrig; 5=sehr hoch);
t-Tests: VA/BDM (t=-0,102; p=0,919);

Tabelle 4: Fragen zum Verständnis der Methoden und strategischen Verhalten

4.3 Beurteilung der Reliabilität

Im Rahmen der Reliabilitätsanalyse wurden zunächst jeweils die Zahlungsbereitschaften der drei Tagessubsamples der drei Methoden auf signifikante Unterschiede untersucht (vgl.

Tabelle 5). Für die Vickrey Auktion sind die Gebote des ersten Erhebungstages signifikant höher als für den zweiten Tag (p=0,023). Dies spricht gegen eine reliable Messung.

Unterschiede zwischen dem ersten und dritten bzw. zweiten und dritten Tag erweisen sich jedoch als nicht signifikant (p=0,329 bzw. p=0,361). Für die BDM und die CVM gibt es

durchgängig keine signifikanten Unterschiede (bei ersterer Methode beträgt die maximale

Differenz der Mittelwerte der drei Tagessubsamples sogar weniger als 50 cent). Generell kann

festgehalten werden, dass der Zufallsfehler bei allen drei Methoden (mit leichtem Vorzug der BDM) als hinreichend gering betrachtet werden kann.

Tagessubsamples Gebote	Vickrey Auktion ^a (n=109)	BDM ^b (n=110)	CVM ^c (n=104)
Tag 1 (MW, SA)	102,46 (60,52)	72,97 (46,28)	174,65 (115,51)
Min, Max	5, 250	10, 200	5, 450
Tag 2 (MW, SA)	70,15 (57,23)	73,08 (62,69)	173,44 (108,41)
Min, Max	1, 200	10, 260	20, 500
Tag 3 (MW, SA)	85,01 (83,10)	72,75 (66,16)	134,32 (81,27)
Min, Max	0,50, 350	15, 250	20, 300

^a t-Tests: Tag1/Tag2 (t=2,321; **p=0,023**), Tag1/Tag3 (t=0,983; p=0,329), Tag2/Tag3 (t=-0,920; p=0,361)

^b t-Tests: Tag1/Tag2 (t=-0,009; p=0,993), Tag1/Tag3 (t=0,015; p=0,988), Tag2/Tag3 (t=0,022; p=0,983)

^c t-Tests: Tag1/Tag2 (t=0,047; p=0,962), Tag1/Tag3 (t=1,493; p=0,141), Tag2/Tag3 (t=1,574; p=0,120)

Tabelle 5: Reliabilitätsanalyse

Zusätzlich zur Analyse der drei Tagessubsamples dienen die zwei Items „Hätte ich noch einmal die Chance ein Gebot abzugeben, würde ich es exakt in der gleichen Höhe abgeben“ bzw. „Hätte ich mir länger über die Höhe des Gebots Gedanken gemacht, hätte ich mich vielleicht anders entschieden“ zur Beurteilung der Reliabilität (vgl. Tabelle 6). Beide Items erhalten für alle drei Methoden eine hohe Zustimmung (Mittelwerte zwischen 3,20 und 4,20 jeweils auf einer 5-stufigen Skala). Dies spricht für eine verlässliche Messung der wahren Zahlungsbereitschaft. Im Detail ist die Zustimmung im Rahmen der BDM am höchsten, für die Vickrey Auktion am niedrigsten. Signifikante Unterschiede gibt es zwischen Vickrey Auktion und BDM (Item 7) und zwischen Vickrey Auktion und BDM bzw. CVM und BDM (Item 8).

Items zur Beurteilung der Reliabilität	Vickrey Auktion (n=109)	BDM (n=110)	CVM (n=104)
Item 7: Hätte ich noch einmal die Chance ein Gebot abzugeben, würde ich es exakt in der gleichen Höhe abgeben ^a	3,84(1,20)	4,20(1,22)	4,02(1,11)
Item 8: Hätte ich mir länger über die Höhe des Gebots Gedanken gemacht, hätte ich mich vielleicht anders entschieden ^b	3,20(1,53)	3,83(1,42)	3,31(1,48)

^a Gemessen auf einer Skala von 1 bis 5 (1=Stimme gar nicht zu; 5=Stimme völlig zu); t-Test: VA/BDM (t=-2,191; **p=0,030**); VA/CVM(t=-1,105; p=0,270); BDM/CVM(t=1,141; p=0,255)

^b Gemessen auf einer Skala von 1 bis 5 (1=Stimme völlig zu; 5=Stimme gar nicht zu; *umkodiert*); t-Test: VA/BDM (t=-2,837; **p=0,005**); VA/CVM(t=-0,448; p=0,654); BDM/CVM(t=2,361; **p=0,019**)

Tabelle 6: Items zur Beurteilung der Reliabilität

Insgesamt scheint die Messung der BDM demnach am verlässlichsten zu sein, die wahre Zahlungsbereitschaft reliabel abzubilden.

4.4 Beurteilung der Face Validität

Zur Überprüfung der Face Validität wurde sowohl das Produktinteresse als auch die Kaufwahrscheinlichkeit der Probanden jeweils mit dem abgegebenen Gebot korreliert. Eine Face-valide Messung fordert diesbezüglich signifikante und positive Korrelationskoeffizienten (je höher das Produktinteresse bzw. die Kaufwahrscheinlichkeit desto höher die Zahlungsbereitschaft). Die Ergebnisse sind in Tabelle 7 zusammengefasst. Die Korrelationsanalysen liefern durchgängig die erwarteten positiven Vorzeichen. Hinsichtlich des Signifikanzniveaus zeigen sich jedoch klare Unterschiede. Während die Vickrey Auktion und die BDM sehr ähnliche Koeffizienten auf hohem Signifikanzniveau zeigen ($r > 0.3$; $p < 0,01$ bzw. $< 0,001$) korreliert das Produktinteresse mit der Zahlungsbereitschaft im Rahmen der CVM deutlich schwächer ($r < 0.2$, $p < 0.05$). Kein signifikanter Zusammenhang lässt sich zwischen CVM Zahlungsbereitschaften und Kaufwahrscheinlichkeit feststellen ($p > 0.05$). Dies könnte auf die fehlende Kaufverpflichtung der CVM zurückzuführen sein.

Unterschiedsanalysen der Korrelationskoeffizienten zwischen den drei Methoden (t-Tests der z-scores) zeigen lediglich signifikante Unterschiede zwischen Vickrey Auktion und BDM bzw. Vickrey Auktion und CVM hinsichtlich Produktinteresse. Die Vickrey Auktion zeigt demnach das stärkste Maß an Face Validität gefolgt von der BDM. Die CVM weist nach dieser Analyse ein deutlich niedrigeres Maß an Face-Validität auf.

Korrelation ^a zwischen Zahlungsbereitschaft und...	Vickrey Auktion (<i>n</i> =109)	BDM (<i>n</i> =110)	CVM (<i>n</i> =104)
Item 9: Wie hoch ist Ihr Interesse an einem iPod? ^b	0.496***	0.368***	0,195*
Item 10: Wie wahrscheinlich ist es, dass Sie in den nächsten 6 Monaten einen iPod kaufen werden? ^c	0.309**	0.373***	0,158 n.s.

^a nach Pearson, Signifikanz einseitig * $p < 0.05$, ** $p < 0.01$, *** $p < 0,001$

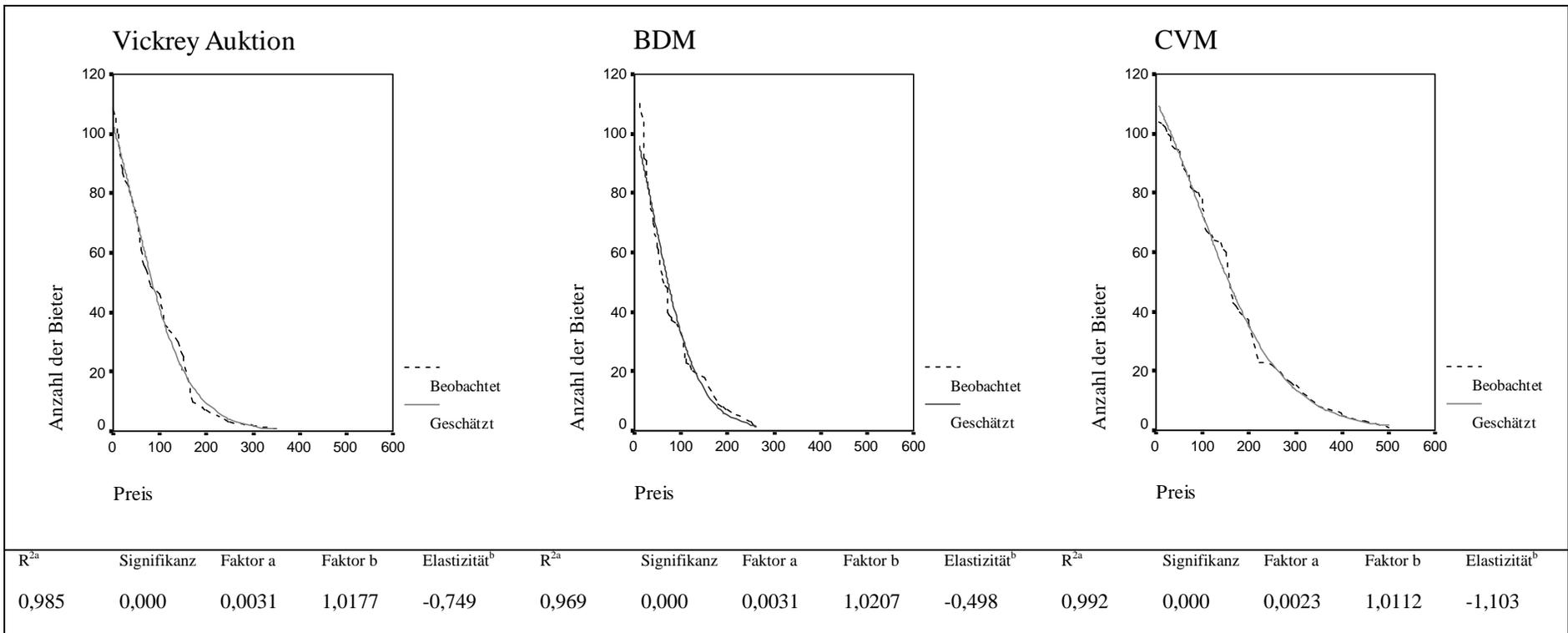
^b t-Tests: VA/BDM ($t=1,656$; $p=0,099$), VA/CVM ($t=2,816$; $p=0,005$), BDM/CVM ($t=1,020$; $p=0,309$)

^c t-Tests: VA/BDM ($t=0,674$; $p=0,501$), VA/CVM ($t=1,534$; $p=0,127$), BDM/CVM ($t=1,012$; $p=0,313$)

Tabelle 7: Face Validität

4.5 Beurteilung der internen Validität

Zur Bestimmung der internen Validität wurde eine Logit-Regression mit der Anzahl der Käufer als abhängige und dem Gebot als unabhängige Variable gerechnet ($y=1/(1/u+a*b^X)$; upper bound, $u=150$). Abbildung 1 zeigt die beobachtete und die vorhergesagte Nachfragekurve. Hohe Fitwerte würden bedeuten, dass die empirisch erhobenen Zahlungsbereitschaften gut durch die geschätzten Absatzreaktionsfunktionen abgebildet werden könnten. Dies gilt als Indikator zur Beurteilung der internen Validität (vgl. Skiera/Revenstorff (1999), Wertenbroch/Skiera (2002)). Die hohen Fitwerte ($R^2 > 0.97$) bestätigen die gute Anpassung und lassen bei allen drei Methoden auf ein zufriedenstellendes Maß an interner Validität schließen. Hinsichtlich R^2 gibt es keine signifikanten Unterschiede zwischen den drei Methoden (Überprüfung mit Hilfe z-transformierter Werte). Darüber hinaus schwanken die Preiselastizitäten der drei Funktionen um den Wert -1 und befinden sich somit im plausiblen Bereich (Tellis (1988)).



^a t-Tests: VA/BDM ($t=0,011$, $p=0,991$); BDM/CVM ($t=0,88$, $p=0,930$), VA/CVM ($t=0,105$, $p=0,917$)

^b berechnet an der Stelle des Medians

Abb. 1: Beobachtete und geschätzte Nachfrage

4.6 Beurteilung der externen Validität

Um die externe Validität der Untersuchung zu überprüfen, wurden den Probanden drei Fragen zur Selbstvalidierung gestellt (7-stufige Ratingskalen von -3 bis +3, vgl. Tabelle 8). Werte nahe 0 weisen auf eine hohe externe Validität hin. Generell liegen die Mittelwerte für alle drei Items und alle drei Methoden im Bereich zwischen -1 und +1 und lassen somit auf ein zufriedenstellendes Maß externer Validität schließen. Im Detail schneidet hier die Vickrey Auktion am besten ab – gefolgt von der BDM. Während sowohl Probanden der Vickrey Auktion als auch jene der BDM tendenziell mehr für den iPod in einem Geschäft zahlen würden als im Rahmen dieser Untersuchung, würden Probanden der CVM tendenziell real weniger bezahlen (Unterschiede zwischen den Gruppen erweisen sich mit Ausnahme vom zweiten und dritten Item zwischen Vickrey Auktion und BDM durchgängig als signifikant).

Selbstvalidierung der Probanden	Vickrey Auktion (n=109)	BDM (n=110)	CVM (n=104)
Item 11: Wenn ich das Produkt jetzt sofort bekäme, würde ich mehr/weniger zahlen ^a	-0,14(1,07)	0,16(0,93)	-0,39(1,16)
Item 12: Im Apple Onlineshop würde ich mehr/weniger zahlen ^b	0,16(1,50)	0,20(1,23)	-0,39(1,16)
Item 13: In einem herkömmlichen Geschäft würde ich mehr/weniger bezahlen ^c	0,03(1,26)	0,57(1,42)	-0,48(1,30)

^a Gemessen auf einer Skala von -3 bis 3 (viel weniger bis viel mehr).

t-Tests: VA/BDM (t=-2,224; **p=0,027**); VA/CVM (t=1,681 **p=0,094**); BDM/CVM (t=3,883; **p=0,000**)

^b Gemessen auf einer Skala von -3 bis 3 (viel weniger bis viel mehr).

t-Tests: VA/BDM (t=-0,237; p=0,813); VA/CVM (t=2,978; **p=0,003**); BDM/CVM (t=3,626; **p=0,000**)

^c Gemessen auf einer Skala von -3 bis 3 (viel weniger bis viel mehr).

t-Tests: VA/BDM (t=0,497; p=0,620); VA/CVM (t=2,873; **p=0,004**); BDM/CVM (t=2,666; **p=0,008**)

Tabelle 8: Selbstvalidierungsfragen zur Beurteilung der externen Validität

5. Zusammenfassung und Ausblick

Ziel dieses Beitrages war der empirische Vergleich zwischen den Methoden Vickrey Auktion und BDM zur Messung von Zahlungsbereitschaften. Aus theoretischer Sicht weisen beide Methoden der realen Kaufangebote den Vorteil der Anreizkompatibilität auf (Becker/deGroot/Marschak (1964), Vickrey (1962)). Dies bedeutet, dass die Bieter bzw. Probanden die dominante Strategie haben sollten, die wahre (maximale) Zahlungsbereitschaft bekannt zu geben. Wirkt diese Anreizkompatibilität auch in der Empirie, sollten beide Methoden c.p. zu den gleichen Ergebnissen kommen. Zum einen sollte also untersucht werden, ob unter sonst gleichen Umständen, der Einsatz der zwei Methoden zu einem ähnlichen Niveau an durchschnittlicher Zahlungsbereitschaft führt (oder ob die zwei Methoden signifikant unterschiedliche Zahlungsbereitschaften ergeben). Zum anderen sollten die beiden Methoden hinsichtlich Unterschiede in den Problembereichen Verständnis bzw. strategisches Verhalten, Reliabilität und Validität analysiert werden.

Hinsichtlich der durchschnittlichen Zahlungsbereitschaft haben sich keine signifikanten Unterschiede zwischen den Methoden Vickrey Auktion und BDM feststellen lassen. Bei entsprechend großen Stichproben sollte es demnach keine Rolle spielen, welche der beiden Methoden zum Einsatz kommt. Die Implikationen für z.B. einen Produktmanager, der auf Basis der durchschnittlichen Zahlungsbereitschaft der intendierten Zielgruppe den Preis eines neuen Produktes festlegen möchte, bleiben die gleichen. Die hypothetische Kalibrierungsmethode CVM ergab hingegen im Durchschnitt eine doppelt so hohe Zahlungsbereitschaft (hoch signifikante Unterschiede im Vergleich zu beiden realen Kaufangeboten). Dies ist konsistent mit bestehenden Vergleichsstudien zwischen realen und hypothetischen Methoden zur Erfassung von Zahlungsbereitschaften (z.B. List/Gallet (2001), Franke/von Hippel (2002), Sattler/Nitschke (2003)).

Nach einer Prüfung der Qualitätskriterien Verständnis bzw. strategisches Verhalten, Reliabilität und Validität ist keine der beiden Methoden klar gegenüber der anderen zu bevorzugen. Die Mechanismen beider Methoden wurden grundsätzlich gut verstanden. Ein beträchtlicher Teil der Befragten (Vickrey Auktion: 15%; BDM: 23%) hat aber trotz aufwändiger Erklärung die Kontrollfrage zum Verständnis falsch beantwortet. Diese Befunde unterstreichen die Wichtigkeit einer sorgfältigen Erklärung der Funktionsweise vor dem Einsatz der Methode. Nur so kann a priori gewährleistet werden, dass die theoretischen Vorzüge der Methoden empirische Entfaltung finden können.

Reliabilitätsanalysen zeigen, dass der Zufallsfehler bei beiden Messmethoden hinreichend gering ist. Leichten Vorteil hat diesbezüglich die BDM. Hinsichtlich der Face Validität ist dies umgekehrt – beide Methoden liefern zufriedenstellende Ergebnisse (durchgängig besser als für die Kalibrierungsmethode CVM) mit leichtem Vorteil für die Vickrey Auktion. Beiden Methoden ist außerdem ein hohes Maß an interner Validität zuzuschreiben. Die Analyse zur externen Validität zeigt, dass die Vickrey Auktion und die BDM valider sind als die hypothetische CVM (mit leichtem Vorteil für die Vickrey Auktion).

In Summe konnte gezeigt werden, dass sowohl die Vickrey Auktion als auch die BDM ein verlässliches und valides Instrumentarium zur Abbildung von Zahlungsbereitschaften darstellen. Bestehende Arbeiten zu Vickrey Auktion *oder* BDM basieren überwiegend auf Untersuchungen, die niedrigpreisige Objekte verwendeten: z.B. Skiera/Revenstorff (1999) (Telefonwertkarten), Sattler/Nitschke (2003) (Mobilfunktarife), Wertenbroch/Skiera (2002) (Coca Cola, Kuchen, Kugelschreiber), Kealy et al. (1990) (Schokoladeriegel), Noussair et al. (2004) (Lebensmittel wie Orangensaft und Kekse), Rutström (1998) (Trüffelschokolade). Aus praktischer Unternehmenssicht sind es aber oft höherpreisige Produkte, bei welchen preispolitische Entscheidungen besonders kritisch sind. Gelten die Befunde bestehender

Arbeiten auch für diese Preissegmente? Die vorliegende Untersuchung ergänzt hierzu die bestehende Literatur: Die Methoden der realen Kaufangebote scheinen auch für höherpreisige Produkte die Nachfrage reliabel und valide abbilden zu können.

Einzigere Unsicherheitsfaktoren bleiben das Verständnis (Sattler/Nitschke (2003), Wertenbroch/Skiera (2002)) und das mögliche strategische Verhalten (Skiera/Revenstorff (1999), Wertenbroch/Skiera (2002)) als potenzielle Störfaktoren, die die theoretischen Vorteile dieser Methoden (v.a. Anreizkompatibilität) gefährden. Da die vorliegende Stichprobe aus der Grundgesamtheit Studierender gezogen wurde (und damit u.a. überdurchschnittliche Bildung ausweist), könnten diese Effekte in anderen Bevölkerungs- bzw. Kundenschichten noch stärker ausgeprägt sein. Zukünftige Forschungsarbeiten könnten diese Problembereiche tiefergehend untersuchen.

Weitere Studien könnten außerdem die zweite wichtige Einschränkung der vorliegenden Untersuchung adressieren: Die Beurteilung der externen Validität beruht in diesem Beitrag auf Selbsteinschätzung der Probanden. Tatsächliches *Verhalten* in einer *anderen* Situation (z.B. reale Kaufsituation im Geschäft) wäre ein verlässlicherer Indikator. Eine mögliche Vorgehensweise skizzieren diesbezüglich Onwujekwe et al. (2004), indem sie das zugrundeliegende Forschungsobjekt wenige Wochen später denselben Kunden zum Kauf angeboten haben, die zuvor an einem Experiment zur Messung von Zahlungsbereitschaften (CVM) teilgenommen haben. Rückschlüsse auf die Prognosequalität bzw. -kraft ließen sich so gezielt untersuchen. Die Facette der externen Validität scheint aus praktischer Sicht besonders relevant: Können z.B. die Ergebnisse über die Messung von Zahlungsbereitschaften *direkt* für Preissetzungsentscheidungen herangezogen bzw. verwendet werden? Wohl nur, wenn sie neben den anderen Facetten auch ein hohes und verlässliches Maß an externer Validität aufweisen.

Verzeichnis der zitierten Literatur

Becker, G.M., DeGroot, M.H., and Marschak, J. (1964). Measuring Utility by a Single-Response Sequential Method. *Behavioral Science*, 9, 226-232.

Ben-Akiva, M., Bradley, M., Morikawa, M., Benjamin, J., Novak, T.P., Oppewal, H., Rao, V. (1994). Combining Revealed and Stated Preferences Data. *Marketing Letters*, 5, 335-350.

Blumenscheid, K., Johannesson, M., Yokoyama, K.K., and Freeman, P.R. (2001). Hypothetical versus real willingness to pay in the health care sector: results from a field experiment. *Journal of Health Economics*, 20, 441-457.

Carson, R., Wright, J., Carson, N, Alberini, A., and Flores, N. (1995). *A Bibliography of Contingent Valuation Studies and Papers*. NRDA, La Jolla, CA.

Dahan, E., Hauser, J.R. (2002). The Virtual Customer. *Journal of Product Innovation Management*, 19, 332-352.

Dodds, W.B., Monroe, K., and Grewal, D. (1991). Effects of price, brand, and store information on buyers' product evaluations. *Journal of Marketing Research*, 28, 307-319.

Fox, J.J., Hayes, D., and Kliebenstein, J. (1998). CVM-X: Calibrating Contingent Values with Experimental Auction Markets. *American Journal of Agricultural Economics*, 80, 455-465.

Franke, Nikolaus, Piller, Frank: Value Creation by Toolkits for User Innovation and Design: The Case of the Watch Market. *Journal of Product Innovation Management*, 21/6, 401-415, 2004

Glass, G.V. and Stanley, J.C. (1970). *Statistical Methods in Education and Psychology*. Englewood Cliffs, NJ: Prentice Hall.

Green, P.E. and Srinivasan, V. (1990). Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice. *Journal of Marketing*, 54/10, 3-19.

Green, P.E., Tull, D.S., and Albaum, G. (1988). *Research for Marketing Decisions*. Englewood Cliffs, NJ: Prentice Hall.

Hoffman, E., Menkhaus, D.J., Chakravarti, D., Field, R.A., and Whipple, G.D. (1993). *Marketing Science*, 12/3, 318-338.

Jorgensen, B.S., Syme, G.J., Smith, L.M., and Bishop, B.J. (2004). Random error in willingness to pay measurement: A multiple indicators, latent variable approach to the reliability of contingent values. *Journal of Economic Psychology*, 25, 41-59.

Kagel, J.H. (1995). Auctions: A Survey of Experimental Research. In: J.H. Kagel, and A.E. Roth (Eds.) *The Handbook of Experimental Economics* (pp.501-581), NJ: Princeton University Press.

Kealy, M.J., Montgomery, M., and Dovidio, J.F. (1990). Reliability and predictive validity of contingent values: Does the nature of the good matter. *Journal of Environmental Economics and Management*, 19, 244-263.

List, J.A. and Shogren, J. (1998). Calibration of the Differences between Actual and Hypothetical Valuations in a Field Experiment. *Journal of Economic Behavior and Organization*, 37/2, 193-205.

List, J.A. and Gallet, C.A. (2001). What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values? *Environmental and Resource Economics*, 20, 241-254.

Loureiro, M.L., McCluskey, J.J., and Mittelhammer, R.C. (2002). Will consumers pay a premium for eco-labeled apples? *The Journal of Consumer Affairs*, 36/2, 203-219.

Maxwell, S. (2002). Rule-based price fairness and its effect on willingness to purchase. *Journal of Economic Psychology*, 23, 191-212.

Mitchell, R.C. and Carson, R.T. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Baltimore, MD: John Hopkins University Press.

Monroe, K.B. (2003). *Pricing: Making Profitable Decisions* (3. Auflage). New York: McGraw-Hill.

Neill, H., Cummings, R., Ganderton, P., Harrison, G., and McGuckin, T. (1994). Hypothetical Surveys and Real Economic Commitments. *Land Economics*, 70/2, 145-154.

Noble, P.M. and Gruca, T.S. (1999). Industrial Pricing: Theory and Managerial Practice. *Marketing Science*, 18/3, 435-454.

Noussair, C., Robin, S., and Ruffieux, B. (2004). Revealing Consumers' Willingness-to-pay: A Comparison of the BDM Mechanism and the Vickrey Auction. *Journal of Economic Psychology*, 25/6, 725-741.

Onwujekwe, O., Hanson, K., Fox-Rushby, J., (2004). Do divergences between stated and actual willingness to pay signify the existence of bias in contingent valuation surveys? *Social Science and Medicine*, 60, 525-536.

Roe, B., Teisl, M.F., Levy, A., and Russel, M. (2001). US consumers' willingness to pay for green electricity. *Energy Policy*, 29, 917-925.

Rowlands, I.H., Parker, P., and Scott, D. (2002). Consumer perceptions of green power. *Journal of Consumer Marketing*, 19/2, 112-129.

Rutström, E.E. (1998). Home-grown values and incentive compatible auction design. *International Journal of Game Theory*, 27, 427-441.

Ryan, M. and San Miguel, F. (2000). Testing for consistency in willingness to pay experiments. *Journal of Economic Psychology*, 21, 305-317.

Sattler, H. and Nitschke, T. (2003). Ein empirischer Vergleich von Instrumenten zur Erhebung von Zahlungsbereitschaften. *Zfbf*, 55/6, 364-381.

Schreier, M. (2005). The value increment of mass-customized products: An empirical assessment and conceptual analysis of its explanation. *Journal of Consumer Behavior*, *forthcoming*.

Schreier, M. and Franke, N. (2005). Tom Sawyer's great law in action: Testing why users are willing to pay more for products they design for themselves. Working Paper. Vienna University of Economics and BA.

Shogren, J.F., List, J.A., and Hayes D.J., (2000). Preference Learning in Consecutive Experimental Auctions. *American Journal of Agricultural Economics*, 82/4, 1016-1021

Simon, H. (1992). *Preismanagement: Analyse, Strategie, Umsetzung*, 2. Auflage. Wiesbaden: Gabler.

Skiera, B. and Revenstorff, I. (1999). Auktionen als Instrument zur Erhebung von Zahlungsbereitschaften. *Zfbf*, 51/3, 224-242.

Skuras, D. and Vakrou, A. (2002). Consumers' willingness to pay for origin labeled wine. *British Food Journal*, 104, 10/11, 898-912.

Vickrey, W. (1961). Counterspeculation, Auctions and Competitive Sealed Tenders. *Journal of Finance*, 16-8-17.

Wertenbroch, K. and Skiera, B. (2002). Measuring Consumer's Willingness to Pay at the Point of Purchase. *Journal of Marketing Research*, 39/2, 228-241.

Wolfstetter, E. (1996). Auctions: An Introduction. *Journal of Economic Surveys*, 10, 367-420.