

Incentivierungsstrategien bei Minderheitenangehörigen

*Ergebnisse eines
Methodenexperiments*

Incentive Strategies for Minorities

Results of an experiment

Patrick Fick und Claudia Diehl

Zusammenfassung

Schriftliche Befragungen stellen unter bestimmten Bedingungen eine kostengünstige Form der Datenerhebung unter Angehörigen ethnischer Minderheiten dar, die ansonsten finanziell sehr aufwändig sind. Ein Vorteil ist, dass eingesparte Mittel in gezielte Maßnahmen zur Steigerung der Ausschöpfungsquote investiert werden können. Hiervon ausgehend werden in dem Beitrag unterschiedliche Formen der Incentivierung und deren generelle und subgruppenspezifische Wirkung auf die Teilnahmebereitschaft und die Stichprobenqualität untersucht. Zunächst werden allgemeine theoretische Überlegungen zum Einsatz von Incentives vorgestellt. Darauf aufbauend wird eine bislang wenig diskutierte „doppelte Incentivierung“ mit einem unbedingten und einem konditionalen Anteil vorgeschlagen. Sie zielt darauf ab, gleichermaßen die Kosten der Nicht-Teilnahme und den Nutzen der Teilnahme für die Befragten zu erhöhen. In einem Methodenexperiment unter deutsch-türkischen jungen Erwachsenen werden diese theoretischen Überle-

Abstract

Mail surveys are under certain conditions a cost-effective mode of data collection in the generally cost-intensive research on ethnic minorities. Compared to other survey modes they make financial resources available that can be invested into strategies to increase survey response rates. Based on these assumptions, we study general as well as sub-group specific effects of various forms of incentives on response rates and sampling bias. We start out by presenting theoretical considerations about the use of different forms of incentives. It follows an introduction of a hitherto little-discussed „double incentive“ strategy consisting of an unconditional and a conditional part. The logic of such a strategy is to increase the costs of non-participation and the benefits of participation for participants at the same time. Empirical evidence is presented on the basis of an experiment with a specific population of young German-Turkish adults. After all, the double incentive strategy has positive effects on the response rate and therefore seems to be a reasonable option.



gungen empirisch überprüft. Dabei erweist sich die doppelte Incentivierung als sinnvolle Strategie mit positiven Effekten auf die Ausschöpfungsquote und die Stichprobenqualität. So können durch den Einsatz materieller Anreize offenbar benachteiligte oder – im Fall von Minderheitenangehörigen – weniger integrierte Subgruppen besser erreicht werden als bei einem Verzicht auf Incentives.

In accordance with prior findings incentives seem to reduce sample bias. Obviously, they lead to a decrease in sampling bias of respondents from disadvantaged or less integrated subgroups.

1 Einleitung¹

Die standardisierte Befragung von Personen mit Migrationshintergrund stellt aus verschiedenen Gründen eine Herausforderung dar. Zwar liegt ihr Bevölkerungsanteil bei rund 20 Prozent (Statistisches Bundesamt 2012), die Heterogenität dieser Population macht jedoch häufig subgruppenspezifische Analysen etwa auf Grundlage der Nationalität notwendig, für die ausreichend hohe Fallzahlen benötigt werden. Die Praxis, Migrantinnen und Migranten lediglich entsprechend ihres Anteils an der Wohnbevölkerung „mitzubefragen“, geht zudem oft damit einher, ausschließlich deutschsprachiges Erhebungsmaterial einzusetzen. Personen ohne ausreichende Deutschkenntnisse können bei dieser Vorgehensweise nicht erreicht werden. Um Fragestellungen aus dem Bereich der Integrationsforschung zu bearbeiten, wird daher meist ein *Oversampling* der interessierenden Gruppen oder ein gesondertes Sample von Subgruppen ähnlicher ethnischer oder nationaler Herkunft benötigt.

Im Fall von „Migrantenbefragungen“ ist ein melderegisterbasiertes Sampling anhand der Nationalität mit anschließender *Face-to-Face* Befragung gängige Praxis. Zum Einsatz kommen aber auch onomastische, d. h. namenbasierte Auswahlverfahren zur Stichprobenziehung aus Telefonverzeichnissen (vgl. Haberland et al. 2011), auf deren Grundlage telefonische Interviews durchgeführt werden.² Beide Methoden sind nicht unproblematisch. Das erstgenannte Verfahren schließt Eingebürgerte aus, so dass bei vielen Fragestellungen der Migrationsforschung eine entlang der abhängigen Variablen verzerrte Stichprobe die Validität der Ergebnisse

1 Wir bedanken uns bei Peter Preisendörfer und Jan Mewes sowie den beiden anonymen Gutachtern für die hilfreichen Hinweise zu einer früheren Version des Manuskripts.

2 Bisweilen werden auch andere Verfahren eingesetzt, etwa eine Ziehung aus dem der Forschung nicht grundsätzlich zur Verfügung stehenden Ausländerzentralregister (vgl. Babka von Gostomski/Pupeter 2008) oder ein Random-Walk-Verfahren nach der BIK Methode wie beim Integrationsurvey des BiB (vgl. Haug/Diehl 2005).

beeinträchtigen kann. Selbst wenn bei der Auswahl der Adressen onomastische Verfahren angewendet werden, bleibt das Problem der gerade bei Migranten sehr hohen Kosten persönlicher Befragungen bestehen (vgl. für *Face-to-Face* Befragungen allgemein Dillman et al. 2009). Diese sind u. a. darauf zurückzuführen, dass es bei vielen Gruppen empfehlenswert ist, bilingual-koethnische Interviewer einzusetzen, um die Befragungsfähigkeit insbesondere von weniger integrierten Befragten zu erhöhen und so die Stichprobenqualität zu verbessern (vgl. Blohm/Diehl 2001; Feskens et al. 2006). Da die wenigsten Erhebungsinstitute einen entsprechenden Interviewer-Pool vorhalten, müssen zumindest für größere Erhebungen neue Interviewer rekrutiert werden. *Telefonisch* sind Migrantinnen und Migranten häufig schwer bzw. eher mobil zu erreichen, so dass diese Erhebungsart mit besonders niedrigen Ausschöpfungsquoten einhergeht (vgl. Granato 1999; Feskens et al. 2006). Diese und verwandte Probleme führen dazu, dass in vielen großen Erhebungen keine gesonderte Migrantenstichprobe gezogen wird (z. B. im Beziehungs- und Familienpanel *pairfam*; für eine Vorstudie siehe: Baykara-Krumme 2010).

Vor diesem Hintergrund erscheint es lohnenswert, über die Alternative der schriftlichen Befragung nachzudenken. Denn zumindest für Fragestellungen von begrenzter Komplexität oder für die Befragung bestimmter Subgruppen wie etwa jüngere Minderheitenangehörige, bietet sie augenscheinlich viele Vorteile. Zwar sind auch schriftliche Befragungen von sinkenden Ausschöpfungsquoten betroffen (vgl. Aust/Schröder 2009), sie sind allerdings preisgünstig und ihre Feldzeit ist in der Regel kürzer als die von *Face-to-Face* Erhebungen (vgl. Schnell 2012).³ Im direkten Vergleich zu anderen Erhebungsmodi werden Ressourcen frei, die in die Anfertigung von mehrsprachigen Erhebungsmaterialien sowie Maßnahmen zur Erhöhung der Ausschöpfungsquote investiert werden können, etwa in eine gezielte Incentivierung der Zielpersonen. Durch solche Anreize können möglicherweise gerade ansonsten schwer zu erreichende Subgruppen, z. B. weniger gut integrierte Personen, eher erreicht werden.

- 3 Die Feldzeit telefonischer Interviews von ethnischen Minderheiten kann unter bestimmten Bedingungen kürzer sein als die von schriftlichen Erhebungen. Dabei kann jedoch das (in Deutschland) dazu häufig notwendige onomastische Auswahlverfahren im Zusammenhang mit dem nötigen Screening sowie die Zahl verfügbarer bilingualer Interviewer die Feldzeit beeinflussen. Dies spiegelt sich in den unterschiedlichen Feldzeiten aktuellerer telefonischer Migrantenbefragungen wider. Ebenso wie die Einbürgerungsstudie von Weinmann/Becher/Babka von Gostomski (2012) berichten Haberland et al. (2011) eine Feldzeit von knapp 35 Tagen, die beide damit mehr als doppelt so lang im „Feld“ waren wie die vom Zentrum für Türkeistudien (2012) jährliche durchgeführte Mehrthemenbefragung mit 14 Tagen.

Schriftliche Befragungen, so unser Ausgangsargument, stellen – in Kombination mit einer gezielten Incentivierung – angesichts der enormen Schwierigkeiten und Kosten, die mit alternativen Erhebungsmethoden einhergehen, bei Migrantenbefragungen unter bestimmten Bedingungen eine gute Wahl dar. Wir möchten uns mit der Frage beschäftigen, ob der Einsatz von Incentives tatsächlich zu höheren Beteiligungsraten führt – und ob sich die Qualität der Stichprobe im Vergleich zur Befragung ohne Incentives verändert, möglicherweise sogar verbessert. Dazu werden im Anschluss an eine Diskussion des Forschungsstands zur generellen und subgruppenspezifischen Wirkung von Incentives (Abschnitt 2) theoretische Argumente zur Art der Incentivierung vorgestellt (Abschnitt 3). Diese sind allgemeiner Natur und keineswegs „migrantenspezifisch“. Sie münden in das Argument, dass eine „doppelte Incentivierung“, die einen bedingungslosen und einen an die Teilnahme geknüpften materiellen Anreiz kombiniert, eine effektive und effiziente Incentivierungsstrategie darstellt. Der darauf folgende Abschnitt 4 beschreibt ein Methodenexperiment unter 500 jungen türkischstämmigen Erwachsenen, in dessen Verlauf Anreize in Form von Bargeld bis zu einer Höhe von 15 Euro eingesetzt wurden. In der Analyse der Ergebnisse (Abschnitt 5) werden die verschiedenen Incentivierungsstrategien hinsichtlich ihrer quantitativen (Ausschöpfungsquote) und qualitativen (Stichproben- und Datenqualität) Auswirkungen untersucht. Besonderes Augenmerk gilt der Frage, ob strukturell, sozial und identifikativ weniger integrierte Subgruppen durch materielle Anreize möglicherweise besonders gut erreicht werden.

2 Auswirkungen der Incentivierung auf die generelle und subgruppenspezifische Teilnahmebereitschaft: Ein Überblick

Incentives, d. h. materielle Befragungsanreize, werden schon lange als Methode zur Erhöhung der Beteiligung an Umfragen eingesetzt und untersucht (vgl. Wotruba 1966; Schnell 2012).⁴ Ihr – insbesondere in Nordamerika – verbreiteter Einsatz geht mit einer kaum mehr zu überschauenden Literatur einher, die unterschiedliche

4 In seiner *Tailored-Design-Method* weist auch Dillman (1978; 2009) auf Incentives als eine mögliche Form zur Steigerung der Umfragebeteiligung hin. Daneben beschreibt er weitere Faktoren (z. B. ein Ankündigungsschreiben, die Zahl der Kontaktversuche, die Art und Farbe des Papiers, der Tinte oder des Briefumschlags, die Nennung des Sponsors oder die Verwendung von Briefmarken) die sich auf die Beteiligung auswirken können (vgl. für entsprechende Evidenzen die Metastudie von Edwards et al. 2002).

Formen der Incentivierung sowie deren Effektivität und Effizienz analysiert und beschreibt (vgl. aber den Überblick bei Singer 2002; siehe auch: www.websm.org). Im Folgenden werden zunächst Forschungsergebnisse präsentiert, die grundsätzlich für oder gegen einen Einsatz unterschiedlicher Incentivierungsformen sprechen. Im Anschluss daran werden Befunde zu etwaigen subgruppenspezifischen Auswirkungen dargestellt.

2.1 Incentivierung und generelle Teilnahmebereitschaft

Bereits eine ältere Metastudie von Yu und Cooper (1983) konnte auf die insgesamt positive Wirkung von Incentives als Anreiz zur Teilnahme an standardisierten Befragungen hinweisen, spätere Metastudien bestätigen dies grundsätzlich (vgl. Church 1993; Singer et al. 1999; Jobber/Saunders/Mitchell 2004; Edwards et al. 2002). Yu und Cooper zeigen, dass mit monetären Anreizen *ceteris paribus* höhere Rücklaufquoten zu erwarten sind (51 vs. 35 Prozent bei Verzicht auf Incentives). Auch für den deutschsprachigen Raum liegen einige Studien zur Wirkung (geldnaher) Incentives vor (vgl. Arzheimer/Klein 1998; Harkness et al. 1998; Diekmann/Jann 2001; Berger et al. 2005; Becker/Imhof/Mehlkop 2007; Mehlkop/Becker 2007; Stadtmüller 2009). Diese bestätigen die positiven Effekte auf die Rücklaufquote; beispielhaft genannt sei hier Stadtmüller (2009), der bereits durch eine dem Anschreiben beigefügte Euromünze eine signifikante Steigerung der Rücklaufquote erzielt (43 vs. 30 Prozent). Allerdings kommen in der Praxis verschiedene Incentivierungsstrategien zum Einsatz, für die unterschiedliche Auswirkungen beschrieben werden.

Monetäre vs. nicht-monetäre Incentivierung

In einer Reihe von Studien wird darauf hingewiesen, dass mit Bargeld bzw. geldnahen Incentives eine höhere Rücklaufquote erreicht werden kann als mit nicht-monetären Anreizen. So konnte im Vergleich zur Auszahlung eines US-Dollars mit einer Lotterieteilnahme keine bessere Rücklaufquote erzielt werden als in der nicht-incentivierten Kontrollgruppe (vgl. Paolillo und Lorenzi 1984). Auch Warriner et al. (1996) können diesen Befund bestätigen (siehe auch: Porter/Whitcomb 2003; für gegenteilige Befunde bei Online-Befragungen: Bosnjak/Tuten 2003). Während Lotterielose also keinen Effekt entfalten, bewirken versprochene Spenden für wohltätige Zwecke sogar einen geringeren Rücklauf (vgl. Warriner et al. 1996). Auch Geschenke wie Kugelschreiber oder Schlüsselanhänger tragen, selbst unter Kontrolle der Höhe ihres materiellen Werts, weniger zur Steigerung der Rücklaufquote

bei als monetäre Incentives (vgl. Edwards et al. 2002; Singer et al. 1999). Ryu/Couper/Marans (2005) zeigen, dass in einer schriftlichen Befragung beim Einsatz eines Gutscheins trotz seines mehr als doppelt so hohen Wertes die Beteiligung niedriger ist als bei Einsatz eines monetären Incentives (45 vs. 67 Prozent).

Die deutschsprachige Forschung konzentriert sich folglich auf die Wirkung geldnaher (vgl. Arzheimer/Klein 1998; Harkness et al. 1998; Porst 1999; Diekmann/Jann 2001) oder bar ausgezahlter Incentives (vgl. Berger et al. 2005; Stadtmüller/Porst 2005; Becker/Imhof/Mehlkop 2007; Mehlkop/Becker 2007; Stadtmüller 2009).

Konditionale vs. unbedingte Incentivierung

Die Strategie *unkonditionaler Incentivierung* (UI) zielt darauf ab, die Befragten mit Hilfe eines vorab und ohne weitere Bedingung überreichten Geschenks zur Teilnahme an der Befragung zu motivieren. Die Logik *konditionaler Incentivierung* (KI) besteht hingegen darin, den Nutzen der Teilnahme mit Hilfe der angebotenen Belohnung für eine erfolgreiche Teilnahme zu erhöhen. Yu und Cooper (1983: 39) zeigen, dass – verglichen mit der Situation ohne Incentives – sowohl monetäre unbedingte („prepaid“) als auch konditionale („promised“) Incentives zu signifikant höheren Rücklaufquoten führen. Singer et al. (1999) bestätigen dies: Beide, KI und UI, bewirken höhere Rücklaufquoten; der Effekt bedingungslos gezahlter Incentives ist dabei tendenziell (aber nicht signifikant) stärker. Church (1993) indes unterstreicht, dass der Einsatz monetärer wie nicht-monetärer Anreize dann wirkungslos bleibt, wenn diese nur *konditional* ausgezahlt werden. Die vorliegenden Metastudien kommen somit zu unterschiedlichen Ergebnissen bezüglich des allgemeinen Nutzens der KI gegenüber einem völligen Verzicht auf Incentives. Eindeutig ist hingegen der Befund einer Gegenüberstellung von konditionalen und unbedingten Strategien: Vorab und bedingungslos gemachte (monetäre) „Geschenke“ haben einen stärkeren Effekt auf die Bereitschaft zur Teilnahme an Befragungen als lediglich versprochene Geschenke.

Im deutschsprachigen Raum konnten sowohl Diekmann und Jann (2001) als auch Becker/Imhof/Mehlkop (2007) diesen Befund bestätigen und zudem auf mögliche negative Effekte einer KI hinweisen. So liegt in beiden Experimenten die Ausschöpfungsquote der KI leicht unter derjenigen der nicht-incentivierten Kontrollgruppe (KG).⁵

5 Die Ausschöpfungsquote der nicht-incentivierten KG liegt im Experiment von Diekmann und Jann (2001) bei 77 Prozent, der konditionalen Gruppe bei 74 Prozent und der unbedingten Gruppe bei 81 Prozent (jeweils 10 CHF als Telefonkarte). Becker/Imhof/Mehlkop (2007) erreichen in der KG einen Rücklauf von 39 Prozent, in der kondi-

Wie viel hilft viel? – Hohe vs. niedrige Incentivierung

Yu und Cooper (1983) haben auf einen starken linearen Zusammenhang zwischen Incentivehöhe (von 0,1 bis 50 Dollar) und Teilnahmebereitschaft hingewiesen. Singer et al. (1999), bei deren Studien Incentives zwischen 1 bis 100 Dollar eingesetzt wurden, sowie Jobber/Saunders/Mitchell (2004) (0,1 bis 5 Dollar) kommen zu ähnlichen Ergebnissen. Weniger eindeutig erscheinen diese Befunde, wenn einzelne Methodenexperimente betrachtet werden, welche die Höhe der Incentivierung unter sonst gleichen Bedingungen variieren und so deren Wirkung direkter erfassen.

Trussell und Lavrakas (2004) liefern in einem Experiment mit schrittweiser Erhöhung des Incentives um einen bis maximal zehn Dollar, einen Nachweis für den annähernd linearen Effekt materieller Anreize. Bolstein und James (1990) testeten andererseits vier unterschiedlich hohe un konditionale Incentives (0,25, 0,5, 1 und 2 Dollar) gegen eine nicht incentivierte KG. Ihren Befunden zufolge können sowohl mit einem als auch mit zwei Dollar signifikante Steigerungen der Rücklaufquote gegenüber der Nicht-Incentivierung erreicht werden; dabei macht es keinen messbaren Unterschied, ob nun mit einem oder zwei Dollar „incentiviert“ wurde. Auch eine spätere Studie spricht gegen einen linearen Effekt der Anreizhöhe auf die Teilnahmebereitschaft (vgl. Bolstein/James 1992). Sie erzielten mit einem un konditional versandten Scheck über 20 Dollar die höchste Rücklaufquote (79 vs. 52 Prozent in der KG). Mit den niedrigeren Beträgen (1, 5 und 10 Dollar) wurden dann zwar geringere Rückläufe erreicht (aber immer noch höhere als in der KG), ihre jeweiligen Abstände zueinander sind jedoch klein und nicht überzufällig. Bei einer Verdoppelung der jeweils un konditional gezahlten 20 auf schließlich 40 Dollar stellt sich überraschend eine um 10 Prozentpunkte geringere Rücklaufquote ein. Ein noch höherer konditionaler Anreiz von 50 Dollar führt dazu, dass die Rücklaufquote unter das Niveau der KG fällt (siehe auch: Singer/Van Hoewyk/Maher 2000). Als Gegenbeispiel sei hier allerdings noch auf die Arbeit von Dykema et al. (2011) hingewiesen. Hier kann eine besonders schwer zu befragende Gruppe (Mediziner) durch außerordentlich hohe materielle Anreize sehr gut erreicht werden: Mit konditional angebotenen 50 Dollar nahmen 15 Prozent (vs. 6 Prozent in der KG) und mit 100 Dollar schließlich 25 Prozent der Mediziner an der Befragung teil.

Berger et al. (2005) haben Ergebnisse einer Untersuchung vorgelegt, bei der, abgesehen von einem bislang nicht weiter veröffentlichten Methodenexperiment im Rahmen des SOEP (vgl. Schupp et al. 2010), erstmalig systematisch die Wir-

Fortsetzung Fußnote 5

tionalen Gruppe 34 Prozent und in der un konditionalen Gruppe 63 Prozent (jeweils 10 CHF in bar).

kung unterschiedlich hoher Incentives im deutschsprachigen Raum getestet wird. Mit steigendem Incentive (10 bzw. 20 DM un konditional sowie 50 DM konditional) gehen keine signifikanten Veränderungen der Rücklaufquote einher, tendenziell sinkt diese sogar.

Diese Ergebnisse machen deutlich, dass Incentives offenbar keine lineare Wirkung auf die Teilnahmebereitschaft haben. Berger et al. argumentieren, dass höhere materielle Anreize durchaus höhere Rücklaufquoten bewirken können, allerdings nur bis zu einer bestimmten Grenze. Die Befunde hierzu sind allerdings uneinheitlich. In einigen Studien finden sich sogar negative Auswirkungen besonders hoher konditionaler Incentives.

2.2 Incentivierung und subgruppenspezifische Teilnahmebereitschaft

Die Leverage-Saliency Theorie weist darauf hin, dass sich jeder einzelne Faktor eines Erhebungsdesigns auf das Teilnahmeverhalten bestimmter Gruppen der Zielpopulation unterschiedlich auswirken kann und dass „the activation of the potential leverage depends on whether the attribute is made salient to the sample person during the survey request“ (Groves/Singer/Cornig 2000: 307). Für jegliche Attribute des Erhebungsdesigns und damit auch für Incentives ergibt sich daraus die Konsequenz, dass deren Einsatz „maßgeschneidert“ an die jeweilige Zielpopulation angepasst (vgl. Dillman 1978) und ihr Einsatz getestet werden sollte. Die Frage, ob und durch welche Incentivierungsstrategie die Teilnahmebereitschaft insgesamt erhöht werden kann, ist primär von forschungspraktischer Relevanz. Entscheidend ist die Frage, ob bestimmte gesellschaftliche Subgruppen *unterschiedlich* auf die Incentivierung reagieren. In diesem Fall bestünde die Gefahr eines Stichprobenbias – oder aber die Chance, durch eine gezielt eingesetzte Incentivierung Ausfälle bei ansonsten schwer zur Teilnahme zu bewegendenden Bevölkerungsgruppen zu verringern.

In vielen Studien werden die zu befragenden Zielpersonen zunächst zufällig unterschiedlichen Incentivierungsbedingungen zugewiesen (Bruttostichproben); anschließend werden die jeweiligen tatsächlich befragten Zielpersonen (Nettostichproben) miteinander verglichen. Durch einen solchen Vergleich kann allerdings keine Aussage darüber getroffen werden, welche „Versuchsgruppe“ am wenigsten Verzerrungen gegenüber der jeweiligen Bruttostichprobe aufweist. Dies gilt etwa für die Studie von Stadtmüller (2009), der auf einen geringeren Anteil an Hauptschülern in der Gruppe stößt, die ein Incentive erhalten hat. Welche Incentivierungsstrategie diesbezüglich die besten Ergebnisse liefert, kann nur anhand eines Vergleichs von Brutto- und Nettostichprobe ermittelt werden. Die Befunde

der entsprechenden Studien sind eher uneinheitlich, wie wir im folgenden Abschnitt zeigen werden.

Incentivierung und Teilnahmebereitschaft von ökonomisch Benachteiligten und ethnischen Minderheiten

Singer et al. (1999) weisen darauf hin, dass bei einem Großteil der untersuchten Studien keine Stichprobenverzerrungen von Incentivierungen ausgehen (vgl. auch James/Bolstein 1990; Warriner et al. 1996; Arzheimer/Klein 1998; Shettle/Mooney 1999). In einigen erhöhen allerdings materielle Anreize die Bereitschaft zur Teilnahme gerade solcher Gruppen, die gemeinhin eine geringere Teilnahmebereitschaft aufweisen („e.g., low income or nonwhite race“). Die Qualität der Stichprobe konnte durch Incentives hier also erhöht werden. Wetzels et al. (2008) zeigen in einem Methodenexperiment des niederländischen Labour Force Surveys, dass materielle Anreize keine verzerrenden Auswirkungen auf die sozioökonomische Zusammensetzung der realisierten Stichprobe haben. Bereits in einer älteren Studie hat Goyder (1994) Informationen u. a. über Geschlecht, Familienstand und Vermögen einbeziehen und zeigen können, dass diese Merkmale Einfluss auf die Teilnahmebereitschaft haben und z. B. Arbeitslose seltener an der Befragung teilnahmen. Der verzerrende Effekt des jeweiligen Merkmals weist in den incentivierten Gruppen zwar in die gleiche Richtung, fällt aber in der Gruppe mit einem un konditional gezahlten Dollar (vs. KG sowie 1 bzw. 10 Dollar konditional gezahlt) am schwächsten aus. Incentives konnten mithin den jeweiligen Bias nicht eliminieren, schwächten ihn im Vergleich mit der KG aber ab. In einer australischen Studie können Barón et al. (2009) die Information über das Ausmaß wohlfahrtsstaatlicher Bezüge der zurückliegenden zwölf Jahre auf Haushaltsebene rekonstruieren und diese Information in einem Experiment nutzen. Dabei zeigt sich ein Bias dergestalt, dass Familien, die nie wohlfahrtsstaatliche Leistungen erhalten haben, deutlich häufiger an der Befragung teilnahmen als Familien, die wiederholt Empfänger von Transferleistungen gewesen sind. Durch den Einsatz materieller Incentives konnte dieser Bias reduziert werden. In einer deutschen Incentivierungsstudie werden Kontextinformationen über das durchschnittliche Einkommen des Ortsteils einbezogen, signifikante Einflüsse der Incentivierung auf die Repräsentativität der Stichprobe werden dabei jedoch nicht gemessen (vgl. Arzheimer/Klein 1998).

Eine der wenigen Studien zu Auswirkungen einer Incentivierung bei Migranten verweist darauf, dass in den Niederlanden ethnische Minderheiten aus nicht westlichen Ländern auf den angebotenen Anreiz kaum reagieren (Wetzels et al. 2008). Zwar steigt die Kooperationsrate der Ausländer von 79 Prozent in der KG auf 93 Prozent unter der Bedingung der höchsten un konditionalen Incentivierung

(autochthone Bevölkerung: 74 vs. 84 Prozent). Dieser Unterschied ist bei den Ausländern – vermutlich aufgrund der sehr kleinen Stichprobengröße – jedoch nicht signifikant. So waren in der höchsten Incentivierungsgruppe lediglich 41 ausländische Personen (gegenüber 355 Autochthonen) enthalten. Die Autoren liefern keine weitere Erklärung dieses Befundes, der in einer weiteren Veröffentlichung detaillierter beschrieben wird (vgl. Feskens et al. 2008). Im Anschluss an die Vertrauensforschung, die zeigt, dass Vertrauen mit steigender sozialer Distanz abnimmt (vgl. Goto 1996), könnte eine denkbare Erklärung sein, dass die für die Teilnahmebereitschaft wichtige Reziprozität zwischen Mehrheits- und Minderheitenangehörigen, die sich als Befragende bzw. Befragte gegenüberstehen, schwerer zu aktivieren ist als zwischen gleichen Gruppen.

Insgesamt betrachtet scheinen Befürchtungen, dass materielle Anreize in standardisierten Befragungen die gewonnene Stichprobe negativ verzerren könnten, auf Basis der vorliegenden Forschungsbefunde unbegründet. Wenn Incentives überhaupt einen Effekt auf die Stichprobenqualität haben, dann eher einen positiven im Sinne einer höheren Teilnahmebereitschaft gerade bei solchen Gruppen, die häufig schwer zur Teilnahme an Surveys zu bewegen sind.

3 Warum und wie wirken Incentivierungsstrategien auf die Teilnahmebereitschaft?

Vor dem Hintergrund der teils widersprüchlichen empirischen Befunde zu den Folgen des Incentiveeinsatzes stellt sich die Frage, wie man unterschiedliche Arten der Incentivierung theoretisch begründen kann, und wie eine „theoriegeleitete Incentivierungsstrategie“ aussehen könnte. Dazu lohnt zunächst ein genauerer Blick auf die Frage, was Befragte dazu bewegt, *ohne* jeglichen materiellen Anreiz an Befragungen teilzunehmen, obwohl diese in der Regel auf freiwilliger Basis stattfinden. Dies ist voraussichtlich dann der Fall, wenn der subjektiv erwartbare Nutzen der Teilnahme den der Nicht-Teilnahme übersteigt (vgl. Esser 1986). Dieser wiederum variiert mit dem Glauben an die Wichtigkeit des Befragungsthemas, dem Gefühl einer „ausgewählten“ Gruppe anzugehören, der Verbundenheit mit oder dem Respekt vor der befragenden Person oder der jeweiligen Institution und dem Gefühl, an etwas Sinnstiftendem beteiligt zu sein (vgl. Dillman 1978; Porst 1999).

Zudem wird bereits durch die Kontaktaufnahme eine Austauschbeziehung eröffnet, in der die Reziprozitätsnorm mit ihrem universellen Charakter Wirkung entfaltet (vgl. Gouldner 1960). Die Reziprozitätsthese geht davon aus, dass Individuen bemüht sind, freiwillig erhaltene Vorteile durch eine Gegenleistung zu erwi-

dern. Die hier relativ abstrakte – und daher sehr schwache – Form von Reziprozität wird erzeugt, indem den Befragten die gesamtgesellschaftliche Notwendigkeit vor Augen geführt und unterstrichen wird, „dass zu guter und aktiver Staatsbürgerschaft einer offenen demokratisch verfassten Gesellschaft auch die Bereitschaft zählt, die Forschung zu unterstützen“ (Schupp 2012: 20). „Die Interviewteilnahme stiftet so durch die Erfüllung der [...] Reziprozitätsnorm dem Befragten einen zusätzlichen Nutzen“ (Arzheimer/Klein 1998: 8). Eine reziproke Situation kann in interviewbasierten Erhebungen auch zwischen Interviewern und Teilnehmern entstehen, wenn sich letztere verpflichtet fühlen, den um ein Interview Bittenden zu helfen.

Es gibt nur wenige Möglichkeiten den geringen Nutzen der Interviewteilnahme zu erhöhen, jedenfalls wenn er in der intrinsisch motivierten „Teilnahme um der Sache willen“ besteht und zudem auf dem Gefühl beruht, eine Reziprozitätsnorm zu erfüllen. In schriftlich und persönlich durchgeführten Erhebungen wird zwar schon durch ein Begleitschreiben bzw. eine Vorankündigung versucht, die Befragten von der Wichtigkeit des Themas zu überzeugen und dabei die freiwillige Teilnahme als wichtigen Bestandteil des sozialen Austauschs („Unterstützung der Forschung“) darzustellen. Kostenseitig steht der Teilnahme aber einerseits die Unsicherheit über die Verwendung der preisgegebenen Informationen und andererseits die zu investierende Zeit gegenüber. Der einzig zu veranschlagende Kostenpunkt der Teilnahmeverweigerung bleibt hingegen psychischer Natur und besteht in der Verletzung von „Höflichkeitsnormen“ (Esser 1986: 39).

Die oben dargestellten empirischen Befunde, denen zufolge ohne Incentivierung offenbar ökonomisch bessergestellte gesellschaftliche Subgruppen leichter zu befragen sind, könnten in dieser Anreizstruktur ihre Ursache haben: Es scheint durchaus plausibel, dass gerade die Unterstützung wissenschaftlicher Forschung einen Nutzenterm darstellt, der für diese meist höher gebildeten Gruppen eher wirkt als für Bildungsferne. Die Unsicherheit über die verwendeten Daten könnte bei dieser Gruppe gleichzeitig höher sein.

Unkonditionale Incentivierung

Angesichts der im letzten Absatz dargestellten Anreizstruktur stellt sich die Frage, wie die Wahrscheinlichkeit der Wahl der Handlungsalternative „Teilnahme“ erhöht werden kann, wenn die Möglichkeiten ausgereizt sind, den perzipierten Nutzen der Interviewteilnahme zu beeinflussen. An dieser Stelle setzt die unkonditionale Incentivierung (UI) an, die versucht, die Kosten der Nicht-Teilnahme zu erhöhen. Dabei geht es darum, durch eine vorab und bedingungslos geleistete „Gabe“ die Bindung an die Reziprozitätsnorm und damit an das – ohne Incentivierung schwa-

che – Gefühl der Verpflichtung zur angemessenen Gegenleistung zu verstärken: „Sending the money with the questionnaire is a goodwill gesture that puts the sponsor and questionnaire in a positive light and sets the stage for the respondent to reciprocate with an appropriate gesture of completing the questionnaire“ (Dillman 2007: 168). Das gezahlte Incentive ist dabei keine Kompensation für die investierte Zeit, sondern vielmehr ein symbolischer Akt, der den Nutzen der Teilnahme nicht berührt. Es stellt daher kein direktes „motive or incitement to action“ (Grant/Sugarman 2004: 720) dar, sondern verteuert vielmehr die Alternative „Nicht-Teilnahme“.

Vor diesem Hintergrund erscheint auch der oben dargestellte Befund plausibel, dass ein besonders hohes unkonditionales Incentive zu einer sinkenden Beteiligung führt: Eine allzu große „Gabe“ dürfte als nicht erwidierbar und deshalb unangemessen wahrgenommen werden und nicht als rein symbolischer Akt, der frei von der impliziten Erwartung einer Gegenleistung ist (vgl. Mauss 1968). Durch eine zu große Vorleistung rückt die Frage nach dem Grund für die großzügige Entlohnung in den Vordergrund und erzeugt Reaktanz, eine Abwehrreaktion auf die „Nötigung“ durch die Forschenden (vgl. Becker/Imhof/Mehlhop 2007). Ab einem bestimmten – individuell aber variierenden – Schwellenwert dürfte die Teilnahmebereitschaft also sinken, weil nun der Nutzen der Nicht-Teilnahme steigt (vgl. Esser 1986). Die Nicht-Teilnahme ist wohl die einzig mögliche Handlung, um die Kontrolle über die soziale Situation wiederzugewinnen und auf die „Nötigung“ zu reagieren.⁶ Sie bleibt zudem völlig folgenlos, weil die Teilnahme an der Befragung freiwillig ist und das Incentive bedingungslos versandt wurde.

Im Gegensatz zur Nicht-Incentivierung dürften bei der Strategie der UI also nicht nur diejenigen an der Befragung teilnehmen, die eine intrinsische Motivation zeigen, die Forschung zu unterstützen, und für die durch die Befragungssituation als solche bereits eine reziproke soziale Beziehung hergestellt werden kann. *Zusätzlich* sollten auch diejenigen teilnehmen, bei denen diese Reziprozität erst durch einen geeigneten materiellen Anreiz erzeugt bzw. verstärkt wird und somit die Kosten der Nicht-Teilnahme gesteigert werden. Dies dürfte besonders für diejenigen Befragten relevant sein, die weniger ausgeprägte Affinitäten zur Wissenschaft aufweisen, z. B. weil sie diesem System aufgrund ihrer eigenen Bildungsferne distanziiert gegenüberstehen. Auch Angehörige ethnischer Minderheiten, die sich nur schwach mit dem Zielland identifizieren, könnten die befragenden Wissenschaftler stärker als out-group Mitglieder wahrnehmen (Johnson et al. 2002)

6 Eine noch schärfere Form der Reaktion auf diese „Nötigung“ wäre es, wenn die Teilnahme verweigert und das Incentives zurückgeschickt wird (immerhin zwei Befragte wählten in dem hier vorgestellten Experiment diese Variante).

und daher weniger geneigt sein, auf den impliziten Appell zu reagieren, der nicht-incentivierten Studien zugrunde liegt: dass die Unterstützung der Forschung eine staatsbürgerliche Pflicht darstelle.

Konditionale Incentivierung – Die rationale Wahl?

Die Strategie der UI geht aus einer Effizienzperspektive mit einer doppelten Ressourcenverschwendung einher. Der Sponsor investiert in diejenigen, die auch ohne Incentive an der Befragung teilgenommen hätten – und belohnt selbst die hartnäckigsten Verweigerer. Bereits Goyder wirft die Frage auf, ob bei *Face-to-Face* Erhebungen „the psychological ploy implied by prepayment might seem provocative“ (1994: 360), weshalb für diese Erhebungsform wohl häufig konditionale Incentives eingesetzt werden. Die rational nahe liegende Strategie, die Auszahlung des Incentives an die Teilnahme zu koppeln, scheint jedoch, wie oben gezeigt wurde, nicht die gewünschten Resultate zu bringen. Angesichts der finanziellen Vorteile, die mit der konditionalen Incentivierung (KI) verbunden sind, soll hier noch einmal die Frage beleuchtet werden, wie dies erklärt werden kann, um anschließend eine modifizierte KI vorzustellen, die gleichzeitig effektiv *und* effizient ist.

Die Logik der KI besteht anders als die der UI darin, den Nutzen der Alternative „Teilnahme“ zu erhöhen. Entsprechend muss der materielle Anreiz hinreichend attraktiv sein, um die oben beschriebenen Kosten der Teilnahme zu übertreffen, oder zugespitzt „attractive enough to tempt people to participate in a research study 'against their better judgement'“ (Grant/Sugarman 2004: 734). Bei genauerer Betrachtung wird schnell klar, warum die Strategie der KI wenig geeignet ist, die Kosten-Nutzen-Kalkulation ins Positive zu wenden. So ist die Befragungssituation für die Befragten ohnehin mit Unsicherheit (z. B. in Bezug auf die Sicherheit ihrer Daten) verbunden (vgl. Esser 1986). Das Versprechen eines bestimmten Geldbetrags ist ein weiterer Unsicherheitsfaktor; es besteht nun zusätzliche Unsicherheit darüber, ob das versprochene „Dankeschön“ für die Teilnahme tatsächlich ausgezahlt wird. Auch mit steigendem Incentive lässt sich lediglich der *potentiell* erwartbare Nutzen erhöhen. Die Unsicherheit über die Wahrscheinlichkeit der Zielerreichung kann so aber nicht beseitigt oder auch nur reduziert werden, möglicherweise steigt diese sogar mit der Höhe des angekündigten konditionalen Incentives, weil dieses Versprechen zunehmend unglaubwürdig wirkt.⁷ Dieses Argument trifft allerdings

7 Dykema et al. (2011) erzielen mit einem besonders hohen konditionalen Anreiz allerdings eine deutliche Verbesserung der Rücklaufquote. Bei einem extrem hohen Betrag erscheint die Möglichkeit, hereingelegt zu werden, nicht unbedingt kleiner, der ausgelegte „Köder“ ist jedoch schlicht zu verlockend, um dieses Risiko nicht einzugehen. Der subjektiv erwartbare Nutzen steigt also, jedoch nicht, weil der „p-Wert“, d. h. die

weniger auf *Face-to-Face* Erhebungen zu. Zwar dürfte die physische Präsenz der Interviewer die Unsicherheit bezüglich der Auszahlung des Incentives erheblich reduzieren, die eigentliche Interviewsituation, in der eine fremde Person Geld anbietet, um Einlass bittet und persönliche Informationen erfragen möchte, bleibt jedoch – abhängig vom „Charme“ der Interviewer – mit einiger Unsicherheit behaftet.⁸

Ein weiteres theoretisches Argument gegen die KI ist, dass die Befragten durch die Ankündigung einer Gratifikation als Gegenleistung für die Interviewteilnahme extrinsisch motiviert agieren – was ihre intrinsische Motivation untergraben kann – und rational kalkulieren. Da in den meisten dargestellten Studien die Höhe des eingesetzten konditionalen Incentives eher gering ist, fällt eine solche rationale Abwägung von Aufwand und finanziellem Ertrag aber offenbar meist zuungunsten der Teilnahme aus. Vor diesem Hintergrund lässt sich leicht erklären, warum eine KI in vielen Studien keine oder sogar eine kontraproduktive Wirkung zeigt: So dürften sich sowohl diejenigen an der Befragung beteiligen, die auch ohne Incentive dazu bereit gewesen wären – *abzüglich* derer, deren intrinsische Motivation durch die Ankündigung der Bezahlung negativ beeinflusst wird und *zuzüglich* derjenigen, die den Nutzen der Befragung aufgrund der KI höher einschätzen als deren Kosten.

Dies dürfte aber eine verhältnismäßig kleine Gruppe sein, da erstens auch in *Face-to-Face* realisierten Befragungen die Unsicherheit der Befragungssituation bestehen bleibt und weil zweitens ein KI häufig zu niedrig ist, um als angemessen und attraktiv wahrgenommen zu werden.

Die Logik doppelter Incentives

Angesichts der bislang dargelegten Argumente für und gegen die unterschiedlichen Incentivierungsformen liegt es nahe, über eine Incentivierungsstrategie nachzudenken, die *gleichermaßen* die Kosten der Nicht-Teilnahme und den Nutzen der Teilnahme erhöht (vgl. Abbildung 1). Eine solche Strategie soll im Folgenden unter

Fortsetzung Fußnote 7

subjektiv wahrgenommene Wahrscheinlichkeit, dass die Teilnahme zum gewünschten Ziel führt, ansteigt, sondern weil der Nutzenterm (U) selbst größer wird.

- 8 Vor diesem Hintergrund ist der Befund bei Goyder (1994) von Bedeutung, der die positiven Befunde der UI in schriftlichen Erhebungen zum Anlass nimmt, diese Ergebnisse für persönlich geführte Interviews zu validieren. Gegen eine nicht-incentivierte Kontrollgruppe testet er ein UI von einem Dollar sowie zwei KI in Höhe von 1 bzw. 10 Dollar. Bemerkenswert ist dabei, dass die von den Interviewern angebotenen 10 Dollar zu keiner signifikant besseren Beteiligung im Vergleich zur Kontrollgruppe geführt haben. Lediglich die zuvor versandte Dollarmünze sorgte für eine gut 12 Prozentpunkte signifikant höhere Beteiligung (vgl. auch Singer et al. 1999).

Abbildung 1 Auswirkungen unterschiedlicher Incentivierungsstrategien auf die Anreizstruktur der Interviewteilnahme

	Zusätzlicher Nutzen der Teilnahme...	Zusätzliche Kosten der Nicht-Teilnahme...
	... im Vergleich zur rein intrinsisch motivierten Teilnahme	
Unkonditionale Incentivierung (UI)	0	+
Konditionale Incentivierung (KI)	(+) ¹	0
Doppelte Incentivierung (DI)	+	+

1 Nur bei Vertrauen in die tatsächliche Auszahlung des konditionalen Anreizes

dem Stichwort der „doppelten Incentivierung“ (DI) diskutiert werden. Ziel ist, mit Hilfe des im ersten Schritt ausgezahlten unkonditionalen Anreizes die bindende Kraft der Reziprozität zu nutzen, d. h. die Kosten der Nicht-Teilnahme zu erhöhen. Das wesentliche Merkmal einer reziproken Situation ist der Aufbau von Vertrauen, insofern wirkt die UI hier auch als ein „symbol of trust“ (Dillman 1978 zitiert nach: Arzheimer/Klein 1998: 10). Dies gilt auch für persönlich durchgeführte Befragungen, bei denen der unkonditionale Anreiz vorab und damit vor der Teilnahmeentscheidung mit einem Ankündigungsschreiben versandt wird und die Kontaktaufnahme durch die Interviewer für die Befragten so in einem anderen, positiven Kontext stattfindet (vgl. Groves/Singer/Cornig 2000). Die UI schafft damit eine wichtige Grundlage für eine KI, denn das Versprechen einer materiellen Belohnung findet in einer sozialen Situation statt, in der die Befragten Vertrauen aufgebaut haben und außerdem hinreichend sicher sein können, dass die Ankündigung auch umgesetzt wird. Dann allerdings sollte ein höheres Incentive eine größere Wirkung zeigen als ein niedriges: Da der konditionale Anteil der DI direkt an eine zuvor erbrachte Leistung gekoppelt ist, sollte diese umso besser wirken, je angemessener und attraktiver der gegebene Anreiz erscheint.

An einer Befragung mit DI Strategie müssten folglich diejenigen – intrinsisch motivierten Befragten – teilnehmen, die auch ohne Incentive teilgenommen hätten, *zuzüglich* derjenigen, bei denen Reziprozität durch ein Incentive hergestellt wird, *zuzüglich* derjenigen, bei denen die Norm der Reziprozität trotz des unkonditionalen Incentives nicht wirkt, die aber auf den zusätzlichen Anreiz reagieren, mit sehr hoher Wahrscheinlichkeit eine (angemessene) Belohnung für ihre Teilnahme zu erhalten.⁹

9 Denkbar wäre zwar auch, dass die intrinsische Motivation einiger Befragter, die „im Dienste der Wissenschaft“ an der Befragung auch ohne jegliche Incentivierung teil

Was die Auswirkungen der unterschiedenen Strategien auf die Teilnahme anbelangt, lassen sich nun folgende Hypothesen formulieren: *Erstens müsste durch die UI die Teilnahme gegenüber der nicht incentivierten, rein intrinsisch motivierten KG steigen (H1)*, da die Reziprozitätsnorm durch eine kleine Vorabgabe gestärkt wird. *Zweitens dürfte eine Erhöhung der UI keine darüber hinausgehenden positiven Auswirkungen auf die Teilnahmebereitschaft zeigen (H2)*, da die Aktivierung der Reziprozitätsnorm keinen Nutzenerwägungen folgt. Diese könnte sogar ausbleiben, wenn die unverlangte Gabe so hoch ist, dass die Befragten glauben, keine Wahl mehr zu haben und mit Reaktanz reagieren.¹⁰ *Drittens dürfte unter der Bedingung einer Kombination aus UI und KI die Teilnahmebereitschaft höher sein als bei der reinen UI (H3)*. Denn durch DI werden nun zusätzlich diejenigen zur Teilnahme bewegt, die weder intrinsisch noch normativ motiviert sind, aber um der angekündigten Belohnung willen teilnehmen, die zu erhalten sie aufgrund der UI als wahrscheinlich einschätzen dürften. *Viertens sollte eine Erhöhung des KI positive Auswirkungen auf die Teilnahmebereitschaft haben (H4)*, da eine nachträglich gezahlte „Belohnung“ umso stärker wirkt, je näher diese an eine angemessene und attraktive Entlohnung des Aufwands heranreicht. *Fünftens erwarten wir, dass durch den Einsatz von Incentives eine Verzerrung zugunsten der ökonomisch bessergestellten Zielpersonen verringert werden kann (H5)*. Diese Hypothese beruht auf dem Argument, dass gerade privilegierte (und häufig höher gebildete) Personen am ehesten ein intrinsisches Interesse an einer wissenschaftlichen Befragung zeigen werden. Und wie oben argumentiert wurde, gehen wir *sechstens* davon aus, *dass Minderheitenangehörige, die starke Bindungen an ihre eigene ethnische Gruppe aufweisen, seltener auf den Appell reagieren, ohne jegliche Incentivierung an einer Befragung teilzunehmen, als andere Minderheitenangehörige (H6)*.

Fortsetzung Fußnote 9

genommen hätten, durch die versprochene Belohnung zerstört wird. Allerdings kann hier spekuliert werden, dass dies nicht unmittelbar geschieht, sondern sich als mittelfristiger Prozess negativ auf die Bereitschaft auswirkt auch ohne Incentivierung an künftigen Befragungen teilzunehmen. Positive Effekte von Incentives in Panelstudien, im Sinne einer geringeren Panelmortalität, sowie ausbleibende Effekte anfänglicher Incentivierung in Folgewellen sprechen allerdings gegen diese Überlegung (vgl. Castiglioni/Pforr/Krieger 2008; Arzheimer/Klein 1998).

10 Allerdings ist fraglich, ob diese Reaktion bereits durch den in der vorliegenden Studie eingesetzten moderaten Incentive ausgelöst wird.

4 Untersuchungsanlage und methodisches Vorgehen

Die Effekte unterschiedlicher Incentivierungsstrategien auf die Ausschöpfungsquote und die Zusammensetzung der realisierten Stichprobe wurde im Rahmen der Pilotstudie „Deutsche/r bleiben?“ untersucht. Die Pilotierung verfolgte das Ziel, die Befragungsbereitschaft junger Erwachsener mit doppelter (deutscher und türkischer) Staatsbürgerschaft zu ermitteln. Aufgrund der geringen Größe dieser Gruppe war eine hohe Beteiligung besonders wichtig, daher wurden verschiedene Maßnahmen zur Steigerung der Ausschöpfungsquote getestet.

Stichprobe und Erhebungsinstrument

Die aus inhaltlichen Gründen interessierende „seltene Population“ besteht aus türkischstämmigen „Optionspflichtigen“, die sowohl die deutsche als auch die türkische Staatsbürgerschaft besitzen (vgl. Diehl/Fick 2012).¹¹ Mit einer registerbasierten Zufallsstichprobe wurden die Adressen von 500 Personen der Jahrgänge 1990 bis 1993 gezogen, dazu wurde auf die Melderegister der Städte Augsburg, Bielefeld, Bremen, Bremerhaven, Dortmund, Düsseldorf und Hannover zurückgegriffen. Neben den bekannten „klassischen“ Maßnahmen zur Erhöhung der Ausschöpfungsquote (vgl. Dillman/Smyth/Christian 2009; Schnell 2012) wurde zumindest ansatzweise auf ein Mixed-Mode-Design zurückgegriffen: Dahinter steht die Idee, die Befragungsteilnahme so angenehm wie möglich zu gestalten, indem der Befragungsmodus frei gewählt werden kann (vgl. Millar/Dillman 2011). Da nur postalische Adressinformationen und keine Telefonnummern der Zielpersonen vorlagen, fanden sowohl der Erstkontakt als auch die Erinnerungsschreiben schriftlich statt, die Teilnahme selbst konnte schriftlich, telefonisch oder via Internet erfolgen, wobei die überwiegende Mehrheit den beigelegten schriftlichen Fragebogen von ca. 15 Minuten Länge nutzte.¹²

11 Optionspflichtige gehören zur zweiten Migrantengeneration, sind also in Deutschland geboren. Daher wurde auf den Versand zweisprachiger Erhebungsmaterialien verzichtet, weil sprachliche Schwierigkeiten bei der Beantwortung nicht zu erwarten waren. Der Methodenbericht der Einbürgerungsstudie von Weinmann/Becher/Babka von Gos-tomski (2012) bestätigt dies; so musste keines der Face-to-Face durchgeführten Interviews mit den türkischstämmigen Optionspflichtigen aufgrund von Sprachproblemen abgebrochen werden.

12 Da mehr als 75 Prozent der Interviews schriftlich erfolgten und keine signifikanten Unterschiede hinsichtlich der Incentivierungsbedingung bestehen, wird auf eine Differenzierung in der Ergebnisdarstellung verzichtet.

Experimentelles Design und Ablauf der Befragung

Zum Start der Befragung (Juni 2011) erhielten alle Befragten ein erstes postalisches Anschreiben.¹³ Die Ausgangsstichprobe wurde dabei zufällig in fünf Subgruppen à 100 Zielpersonen eingeteilt, bestehend aus einer Kontrollgruppe sowie vier unterschiedlichen Incentivegruppen. Aus pragmatischen Gründen und aufgrund der geschilderten recht eindeutigen Befunde existierender Studien wurde dabei auf eine rein konditionale Incentivierung verzichtet.

- I. Der *Kontrollgruppe* (KG) wurde kein materieller Teilnahmeanreiz gewährt. Wie in den übrigen Gruppen auch, lag dem Erhebungsmaterial ein Anschreiben bei, welches mit folgendem Satz abschloss: „Wir bedanken uns im Voraus für Ihre Mitarbeit und Ihren Beitrag zur sozialwissenschaftlichen Forschung.“
- II. Die zweite Gruppe erhielt ein *unkonditionales Incentive in Höhe von 5 EUR* (UI 5). Dazu wurde in der unteren freien Ecke des Anschreibens ein Fünfeuroschein aufgeklebt, darunter der Hinweis „5 € für Ihre Mühe als kleines Dankeschön!“.
- III. Um die Wirkung der Höhe des Anreizes messen zu können, erhielt die dritte Gruppe ein *unkonditionales Incentive* in Höhe von 10 EUR (UI 10).
- IV. Schließlich sollte die Strategie *doppelter Incentivierung mit 5 + 5 EUR* (DI 5+5) getestet werden. Analog zur Gruppe zwei war hier ein Fünfeuroschein als Dankeschön beigelegt. Abweichend von dem sonst völlig identischen Anschreiben wurde das Anschreiben allerdings um den folgenden – optisch hervorgehobenen – Satz ergänzt: „Sollten Sie an unserer Befragung teilnehmen, werden wir Ihnen zusätzlich zu dem beigelegten Geschenk eine Aufwandsentschädigung in Höhe von 5 € zukommen lassen.“
- V. Mit der Strategie der *doppelten Incentivierung mit 5 + 10 EUR* (DI 5+10) wurde die Wirkung der Höhe des versprochenen Anreizes getestet. Analog zur vierten Gruppe wurden hier zehn statt fünf Euro versprochen. Ausgehend von einer Interviewdauer von unter einer halben Stunde können 10 EUR gerade bei der jungen Zielgruppe als eine angemessene Entlohnung betrachtet werden.

Sieben Tage nach dem Start der Befragung erfolgte eine erste Nachfassaktion mittels Postkarte. Die zweite Nachfassaktion fand drei Wochen nach dem Start statt

13 Diese erste Aussendung bestand aus einem Anschreiben, in dem auf Datenschutzaspekte, die Herkunft der Adressdaten sowie die Freiwilligkeit der Teilnahme hingewiesen wurde, einem selbstauszufüllenden Fragebogen inkl. eines freigemachten Rückumschlags, einer Postkarte zur Übermittlung einer Telefonnummer und etwaiger Rückruftermine sowie dem Link zur Onlineumfrage inkl. Kennwort. Um die Teilnahme möglichst interessant zu gestalten, wurde das Anschreiben mit einer persönlichen Anrede, einem farbigen Briefkopf der Georg-August-Universität Göttingen sowie einer sich ebenfalls farblich abhebenden Unterschrift versehen.

und bestand aus einer Sendung, die identisch mit der ersten war. Während das bereits ausgezahlte Incentive in den Nachfassaktionen gegenüber keiner der Gruppen Erwähnung fand, wurden die beiden doppelt incentivierten Gruppen in beiden Nachfassaktionen an die versprochene „Belohnung“ für die Teilnahme mit einer Hervorhebung im Text erinnert und so eine entsprechende Salienz des Incentives hergestellt.

Anreicherung der Ausgangsstichprobe mit Geodaten

Die Ausgangsstichprobe wurde mit Geodaten der Firma microm¹⁴ angereichert, um Aussagen darüber machen zu können, ob bestimmte Subgruppen von Zielpersonen durch die unterschiedlichen Incentivierungsbedingungen eher zur Teilnahme motiviert werden können als andere. Auf der Grundlage dieser Daten stehen – wenn auch auf Aggregatebene – jenseits der im Melderegister verfügbaren Angaben Informationen für die gesamte Ausgangsstichprobe und nicht nur für die realisierte Stichprobe zur Verfügung. Damit können mögliche Stichprobenverzerrungen in den unterschiedlichen Incentivegruppen identifiziert werden. Entsprechend unseren theoretischen Argumenten gilt unser Interesse vor allem den ökonomisch schlechter gestellten bzw. weniger gebildeten Zielpersonen sowie denjenigen, die starke Bindungen an ihre ethnische Gruppe aufweisen. Als indirekte Indikatoren für diese beiden Individualmerkmale werden Brutto- und Nettostichprobe im Hinblick auf die lokale *Kaufkraft* und den *Ausländeranteil* miteinander verglichen. Diese zugespielten microm-Indikatoren bilden das nächste Wohnumfeld der Befragten auf der Ebene der Straßenabschnitte ab. Die Straßenabschnittsebene ist eine relativ kleine Untersuchungseinheit, in unserem Datensatz liegen hier pro Individuum aggregierte Informationen von durchschnittlich 130 Privathaushalten vor, die in einem Haus oder einem Straßenabschnitt wohnen.

Der *Kaufkraftindex* bildet die durchschnittliche Kaufkraft des Straßenabschnitts ab, in dem der oder die Befragte wohnt. Der microm Bundesdurchschnitt ist dabei auf den Wert 100 festgesetzt, darunter liegende Werte stehen somit für eine unterdurchschnittliche Kaufkraft des Straßenabschnittes. Der *Ausländeranteil* liegt als ordinale Variable vor (von 0 „niedrigster Anteil“ bis 8 „höchster Anteil“). Auf Straßenabschnittsebene lässt die Variable eine Einschätzung über die Zusammensetzung der Nachbarschaft des Haushaltes zu, dem der oder die jeweilige Befragte angehört.

14 Weitere Informationen über die hauptsächlich im Bereich des Konsumentenmarketings verwendeten Geodaten unter www.microm-online.de. Mit einem ähnlichen Vorgehen wurde der spezielle MICROM-SOEP-Datensatz erstellt (vgl. Goebel et al. 2007).

Für einen Vergleich der unter verschiedenen Incentivierungsbedingungen realisierten Nettostichproben wird schließlich auf Informationen aus der eigentlichen Befragung zurückgegriffen. Zur Untersuchung etwaiger Stichprobenverzerrungen wird hier der *Abiturientenanteil*, der Grad der *Identifikation als Deutsche/* sowie die *Identifikation als Türke/Türkin* herangezogen. Zur Einschätzung der Datenqualität wird der *Item-Nonresponse* gemessen und schließlich auch der Einfluss der Incentivierung auf die *Wiederbefragungsbereitschaft* untersucht. Schließlich werden auch die *Kosten pro realisiertem Interview* ausgewiesen.¹⁵

5 Ergebnisse: Der Einfluss der Incentivierung auf die Teilnahmebereitschaft und die Stichprobenqualität

Die Ergebnisse werden in Abbildung 2 zunächst grafisch, mit Hilfe der aus der Ereignisdatenanalyse bekannten Kaplan-Meier-Methode, in Form der failure function (kumulierte Ratenfunktion) $H(t)$ dargestellt (vgl. Blossfeld/Golsch/Rohwer 2007: 72f.):

$$\hat{H}(t) = -\log(\hat{G}(t))$$

Die failure function ist die komplementäre Überlebensfunktion $G(t)$, beschrieben als:

$$\hat{G}(t) = \prod_{l: \tau_l < t} \left(1 - \frac{E_l}{R_l}\right)$$

mit:

$$\hat{G}(0) = 1$$

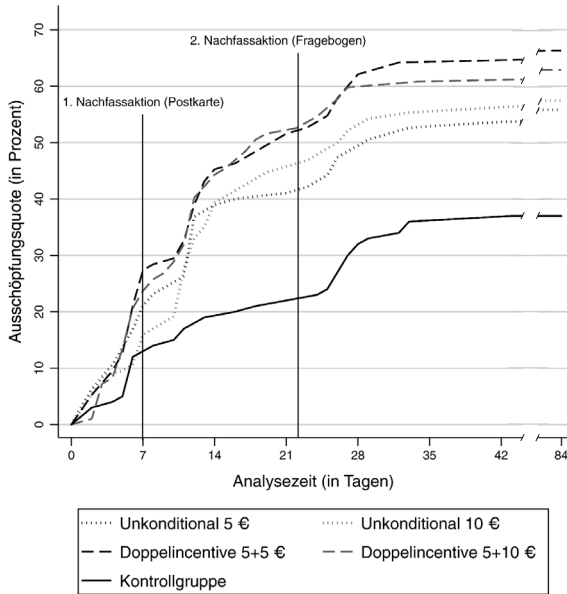
E_l = Zahl der Personen, die zum Zeitpunkt τ_l bereits geantwortet haben

R_l = Zahl der Personen, die zum Zeitpunkt τ_l noch antworten könnten

Dabei können die oben formulierten Erwartungen zum Teil bestätigt werden (vgl. Tabelle 1). Die Teilnahmebereitschaft ist, entsprechend der ersten formulierten Hypothese, bei einer UI deutlich höher als in der Kontrollgruppe, in der kein materieller Anreiz angeboten wurde (37 vs. 55 bzw. 57 Prozent). In der grafischen Darstellung ist zudem auffällig, dass sich die Teilnahmebereitschaft in der KG nach der

15 In die Kostenrechnung sind die Ausgaben für Druck und Versand der bis zu drei Aussendungen, das Incentive inkl. der Versandkosten des konditionalen Incentiveanteils sowie die Portokosten der Fragebogenrücksendung eingegangen. Diese Summe wurde anschließend durch die Anzahl der realisierten Interviews dividiert.

Abbildung 2 Failure-function über die Dauer der Feldphase nach Versuchsgruppen



Quelle: Deutsche/r bleiben? (n=483)

ersten Nachfassaktion deutlich verlangsamt und hier die zweite Nachfassaktion offenbar besonders wichtig war, um zur Teilnahme zu motivieren. Hinsichtlich der zweiten Hypothese zeigt sich, dass es zwischen den beiden vorab incentivierten Gruppen kaum Unterschiede in der Teilnahmebereitschaft in Abhängigkeit von der Höhe des unkonditionalen Anreizes gibt (55 vs. 57 Prozent). Ebenfalls erwartungsgemäß (H3) steigt die Teilnahmebereitschaft bei DI nochmals an (auf 66 Prozent bei DI 5+5). Allerdings führt *entgegen* der oben in H4 formulierten Erwartung eine Erhöhung des konditional ausgezahlten Incentiveanteils von 5 auf 10 EUR nicht zu einer weiteren Steigerung der Teilnahmebereitschaft (66 vs. 63 Prozent). Eine moderate Kombination von unkonditionaler und konditionaler Incentivierung scheint also, zumindest bei der hier untersuchten Gruppe, die besten Resultate zu liefern. Wenngleich die Unterschiede zwischen den vier Incentivierungsgruppen teilweise recht deutlich sind, sind sie – wohl vor allem bedingt durch die geringe Fallzahl – nicht signifikant voneinander verschieden.

Da die entscheidende Erfolgsgröße allerdings in der Regel die Stichprobenqualität und nicht die Ausschöpfungsquote ist, wird in einem zweiten Schritt genauer untersucht, ob bestimmte Subgruppen besonders stark auf eine bestimmte

Tabelle 1 Ausschöpfungsquoten und Kosten pro Interview unter verschiedenen Incentivierungsstrategien (n in Klammern)

	KG	UI 5	UI 10	DI 5+5	DI 5+10	Gesamt
Bruttostichprobe (in %)	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)	100 (500)
Stichprobenneutrale Ausfälle (in %)	0,0 (0)	4,0 (4)	5,0 (5)	5,0 (5)	3,0 (3)	3,4 (17)
Nettostichprobe (in %)	100 (100)	100 (96)	100 (95)	100 (95)	100 (97)	100 (483)
Ausschöpfungsquote (in %) ^{1,2}	37,0 (37)	55,2** (53)	56,8** (54)	66,3** (63)	62,9** (61)	55,5 (268)
Preis pro Interview (in Euro)	14,43	19,07	27,07	21,53	27,04	22,43

1 Ausschöpfungsquote = Interviews / (Versandte Fragebögen – Ausfälle); entspricht RR5 der AAPOR (2011) Standard-Definition

2 Chi-Quadrat-Test (jeweils vs. KG)

** $p < .05$

KG: Kontrollgruppe; UI 5: Unkonditional 5 EUR; UI 10: Unkonditional 10 EUR; DI 5+5: Doppelte Incentivierung 5+5 EUR; DI 5+10: Doppelte Incentivierung 5+10 EUR

Quelle: Deutsche/r bleiben? (n=500)

Incentivierung reagieren. Dazu werden unter Einbezug der zugespielten Makrodaten zunächst Brutto- und Nettostichprobe miteinander verglichen (vgl. Tabelle 2). Die Informationen über beide Stichproben sind naturgemäß begrenzt, daher werden schließlich auch die unterschiedlichen Nettostichproben miteinander verglichen, um Zusammenhänge zwischen der Befragungsteilnahme und individuellen Merkmalen zu untersuchen. Je ähnlicher die jeweiligen Nettostichproben der Kontrollgruppe sind, desto eher kann davon ausgegangen werden, dass durch den Einsatz von Incentives kein bzw. nur ein geringer Stichprobenbias verursacht wurde.

In Tabelle 2 zeigt sich zunächst – entsprechend der in H5 formulierten Erwartung – dass Zielpersonen, die ohne Incentives (KG) um eine Teilnahme gebeten wurden und dieser Bitte nachgekommen sind, tendenziell (wenngleich nicht statistisch signifikant) häufiger in bessergestellten Wohngebieten leben (operationalisiert über den Kaufkraftindex) als diejenigen, deren Teilnahme incentiviert wurde. Bezüglich des Ausländeranteils in der Wohngegend zeigt sich keine systematische Verzerrung der Netto- gegenüber der Bruttostichprobe, H6 kann also nicht bestätigt werden. Der Blick auf den Vergleich des Frauenanteils in Brutto- und Nettostichprobe zeigt zudem, dass der Anteil der teilnehmenden Frauen in der KG erhöht ist: Ohne Incentivierung haben offenbar überdurchschnittlich viele Frauen an der Befragung teilgenommen, die diesbezüglich eher intrinsisch motiviert zu sein scheinen.

Tabelle 2 Vergleich von Brutto- und Nettostichprobe unter verschiedenen Incentivierungsstrategien

	KG	UI 5	UI 10	DI 5+5	DI 5+10	Gesamt
	Mittelwertdifferenzen (= Teilnehmer – Nicht-Teilnehmer) ¹					
Kaufkraftindex (BRD=100) ²	5,2**	0,3**	1,6**	-1,9**	-2,5**	0,3**
Ausländeranteil (Min: 0; Max: 8) ²	-0,1**	0,3**	0,1**	0,2**	-0,2**	0,1**
Frauenanteil ³	23,4**	5,3**	6,8**	0,7**	-2,0**	7,9**

1 Beispiel: Mittelwertdifferenz des Frauenanteils in der KG: 23,4 Prozentpunkte (= Anteil der Teilnehmerinnen: 56,7 % - Anteil der Nicht-Teilnehmerinnen: 33,3 %)

2 zweiseitiger t-Test

3 Chi-Quadrat-Test

** $p < .05$; * $p < .10$

KG: Kontrollgruppe; UI 5: Unkonditional 5 EUR; UI 10: Unkonditional 10 EUR; DI 5+5: Doppelte Incentivierung 5+5 EUR; DI 5+10: Doppelte Incentivierung 5+10 EUR

Quelle: Deutsche/r bleiben? (n=483)

Dies entspricht einem Ergebnis bei Lesser et al. (2001), auch bei ihnen ist das Geschlechterverhältnis nur in den incentivierten Gruppen ausgeglichen. Insgesamt bestätigen unsere Befunde die oben dargestellten Studien insofern als deutlich wird, dass durch eine Incentivierung der Stichprobenbias eher reduziert als vergrößert wird. Dieses Resultat zeigt, wie wichtig ein Vergleich von Brutto- und Nettostichprobe ist: Es kann keinesfalls – wie bei einem Vergleich der realisierten Stichproben untereinander – davon ausgegangen werden, dass die KG die „richtige“ Referenzgröße darstellt.

Ein Nachteil des Brutto-Netto-Vergleichs ist allerdings, dass bei diesem Vorgehen die konstatierten Zusammenhänge nur indirekt untersucht werden können, liegen doch mit den Makroindikatoren Kaufkraftindex und Ausländeranteil in der Nachbarschaft sehr crude Proxies für die theoretisch interessierenden Individualmerkmale vor. Abschließend soll deshalb untersucht werden, ob sich die unter unterschiedlichen Incentivierungsbedingungen realisierten Stichproben im Hinblick auf die theoretisch relevanten Individualmerkmale der Befragten unterscheiden (vgl. Tabelle 3). Auch die Datenqualität, operationalisiert über die Informationen zu Item-Nonresponse, sowie die Bereitschaft für eine erneute Befragung soll in den Blick genommen werden. Häufig wird angenommen, dass Teilnehmerinnen und Teilnehmer einer Befragung „ihre Arbeit“ besser erledigen und sorgfältiger antworten, wenn sie dafür einen Anreiz erhalten haben (vgl. Stadtmüller/Porst 2005).

Der Brutto-Netto-Vergleich hat bereits gezeigt, dass – entsprechend unserer in H5 formulierten Erwartung – Zielpersonen, die nicht incentiviert wurden, in Nachbarschaften mit einem tendenziell höheren Kaufkraftindex leben. Wir waren

Tabelle 3 Ausgewählte Items in der realisierten Stichprobe nach
Versuchsgruppen

	KG	UI 5	UI 10	DI 5+5	DI 5+10	Gesamt
Abiturientenanteil (in %) ¹	43,2	37,7	33,3	22,2**	26,2*	31,3
Identität als Deutsche/r (Min: 0; Max: 4) ²	2,2	1,6**	2,0	1,9	1,9	1,9
Identität als Türke/in (Min: 0; Max: 4) ²	2,7	3,0	3,2*	3,0	2,8	2,9
Erneute Befragung erwünscht (in %) ¹	69,4	76,9	64,2	65,6	83,6	72,2
Item-Nonresponse (Min: 0; Max: 35) ³	1,7	1,1	1,3	1,4	1,2	1,3

1 Chi-Quadrat-Test

2 Tau-b

3 zweiseitiger t-Test

** $p < .05$; * $p < .10$ (jeweils vs. KG)

KG: Kontrollgruppe; UI 5: Unkonditional 5 EUR; UI 10: Unkonditional 10 EUR; DI 5+5: Doppelte Incentivierung 5+5 EUR; DI 5+10: Doppelte Incentivierung 5+10 EUR

Quelle: Deutsche/r bleiben? (n=268)

davon ausgegangen, dass ökonomisch bessergestellte Zielpersonen häufig auch eine höhere Bildung und deshalb eine höhere intrinsische Motivation zur Teilnahme an den Tag legen. Sehr eindrücklich bestätigt dies nun ein Blick auf die Abiturientenquote auch auf der individuellen Ebene. In allen incentivierten Gruppen befinden sich deutlich weniger Personen mit (Fach-)Hochschulreife als in der Kontrollgruppe, deren Anteil ist insbesondere in der mit zwei mal fünf Euro incentivierten Gruppe (DI 5+5) signifikant von der KG verschieden und nur fast halb so hoch. Die „Überrepräsentation“ ökonomisch Bessergestellter in der KG zeigt sich auch bei einem hier nicht ausgewiesenen Vergleich der Kaufkraftindizes in den realisierten Stichproben: So befinden sich in der Gruppe DI 5+5 signifikant mehr Befragte aus ökonomisch benachteiligten Wohngebieten als in der KG.

Darüber hinaus zeigt Tabelle 3 zumindest tendenziell, dass die Befragten, die nicht incentiviert wurden, sich entsprechend unserer Erwartung (H6) weniger stark mit der eigenen ethnischen Gruppe identifizieren als die incentivierten Zielpersonen (beim Vergleich von UI 10 zur KG ist dieser Unterschied signifikant). Umgekehrt ist in der KG die Identifikation als Deutsche/r besonders hoch, was ebenfalls in Richtung des oben formulierten Arguments weist, dass der Appell an die „staatsbürgerschaftliche Pflicht“ zur Unterstützung der Forschung (vgl. Schupp 2012) durch eine Identifikation mit dem Zielland gefördert wird. Angesichts der geringen Unterschiede zwischen den Gruppen und der relativ kleinen Stichprobengröße sollte dieses Ergebnis allerdings nicht überbewertet werden, zumal wir keinen konsistenten Unterschied zwischen den unterschiedlichen Incentivierungsbedingungen feststellen können.

Hinsichtlich der Datenqualität und der Wiederbefragungsbereitschaft bestätigen unsere Befunde existierende Studien (vgl. Singer et al. 1999). Die Auswirkungen auf die Datenqualität sind schwach und eher positiv – im Sinne eines geringeren Item-Nonresponse im Falle der Incentivierung. Hinsichtlich der Wiederbefragungsbereitschaft zeigt sich kein eindeutiger Befund im Vergleich von Kontroll- und Incentivegruppen. Dies entspricht im Kern dem Befund von Singer (1998), ihr zufolge haben Incentives keinen positiven Einfluss auf die *Absicht* an zukünftigen Befragungen teilzunehmen, wohl aber auf das *Handeln*, wenn mit Abstand eines halben Jahres eine erneute Befragung durchgeführt wird (siehe auch: Shettle/Mooney 1999).

Lässt man abschließend die geschilderten qualitativen Aspekte der gewonnenen Stichprobe außer Acht, so ist unter reinen Kostengesichtspunkten die KG mit gut 14 Euro pro erfolgreichem Interview (vgl. Fußnote 15) die günstigste Variante (vgl. Tabelle 1). Bemerkenswert ist jedoch, dass im Vergleich von KG zu UI 5 oder DI 5+5 die Kostendifferenzen geringer sind als die jeweils eingesetzten Incentives. Dies trifft auch auf die Kostendifferenz von UI 5 zu DI 5+5 sowie von DI 5+5 zu DI 5+10 zu. Der schnellere Rücklauf und der damit möglich gewordene Verzicht auf weitere Nachfassaktionen hat sich kostendämpfend ausgewirkt.

6 Zusammenfassung und Ausblick

Je nach interessierender Subgruppe stellt sich in der Integrationsforschung häufig das Problem sehr kleiner Ausgangspopulationen und einem entsprechend kleinen Pool von Adressen oder Telefonnummern (z. B. bei Neuzuwanderern oder Angehörigen der dritten Generation). Darüber hinaus werden für differenzierte Analysen im Bereich der Integrationsforschung ausreichend hohe Fallzahlen für einzelne ethnische Gruppen benötigt. In allgemeinen Bevölkerungsumfragen werden diese Fallzahlen ohne ein *Oversampling* häufig nicht erreicht. Spezifische Migrantenbefragungen sind, gerade wenn sie als *Face-to-Face* Interviews durchgeführt werden, sehr kostspielig. Daher wurde im vorliegenden Beitrag argumentiert, dass eine schriftliche Befragung von Personen mit Migrationshintergrund für bestimmte Fragestellungen eine sinnvolle und günstige Alternative darstellen kann. Die frei werdenden finanziellen Mittel können dann in Maßnahmen zur Erhöhung von Ausschöpfung und Datenqualität investiert werden, beispielsweise in eine gezielte Incentivierungsstrategie.

In diesem Beitrag wurden zunächst theoretische Argumente zusammengetragen, die für den Einsatz einer Kombination von *unkonditionaler* und *konditio-*

nalere Incentivierung sprechen. Eine solche Strategie *doppelter Incentivierung* zielt darauf ab, die Kosten der Nicht-Teilnahme zu erhöhen, indem Reziprozität hergestellt bzw. verstärkt wird. Zusätzlich wird der erwartete Nutzen der Teilnahme gesteigert, indem Vertrauen aufgebaut und den Befragten so eine hinreichende Sicherheit darüber gegeben wird, das versprochene Incentive nach der Teilnahme tatsächlich zu erhalten.

Bei der empirischen Überprüfung dieser Annahmen haben sich die Vorteile einer doppelten Incentivierung in der hier dargestellten Befragung deutsch-türkischer junger Erwachsener zumindest ansatzweise bestätigt. In der doppelt incentivierten Gruppe wurde eine Ausschöpfungsquote von bis zu 66 Prozent erreicht – gegenüber 37 Prozent in der nicht-incentivierten Kontrollgruppe und 55 bzw. 57 Prozent in den rein un konditional incentivierten Gruppen.

Wenngleich die Ergebnisse insgesamt vor dem Hintergrund der geringen Stichprobengröße mit Vorsicht zu interpretieren sind, scheinen Befürchtungen, materielle Anreize in standardisierten Befragungen könnten die gewonnene Stichprobe negativ verzerren oder die Datenqualität mindern, auf Basis der vorgestellten Befunde unbegründet. Wenn Incentives überhaupt einen Effekt auf die Stichprobe und die Qualität der Daten haben, dann eher einen positiven im Sinne einer Verringerung der Nichtteilnahme von „Problemgruppen“ und einer verbesserten Datenqualität: Mit den Geodaten lagen Makroindikatoren vor, die einen Vergleich von Brutto- und Nettostichprobe ermöglichen. Bei diesem Vergleich ist deutlich geworden, dass die Teilnehmerinnen und Teilnehmer der nicht-incentivierten Gruppe tendenziell aus wohlhabenderen Wohngebieten stammen, die, so das theoretische Argument, eine höhere Bildung und eine stärkere intrinsische Motivation zur Beteiligung an wissenschaftlichen Umfragen aufweisen. Der Brutto-Netto-Vergleich hat in Übereinstimmung mit existierenden Befunden gezeigt, dass Frauen signifikant häufiger auch ohne materielle Anreize an der Befragung teilnahmen. Bei dem weitergehenden Vergleich spezifischer Items in den unter unterschiedlichen Incentivierungsbedingungen realisierten Stichproben haben sich diese Befunde insofern bestätigt, als in den doppelt incentivierten Subgruppen weniger hoch Gebildete stärker vertreten sind als in der nicht incentivierten Kontrollgruppe. Im Hinblick auf die (identifikative) Integration zeigt sich dieser Zusammenhang nur tendenziell, zudem scheint diesbezüglich ausschlaggebend zu sein, dass überhaupt incentiviert wird, d. h. es lässt sich kein Unterschied zwischen un konditionaler und doppelter Incentivierung feststellen.

Mit Verweis auf die Leverage-Saliency Theorie wurde verdeutlicht, dass Incentives sowie alle weiteren Merkmale des Erhebungsdesigns auf verschiedene Gruppen unterschiedlich auswirken können, insofern können die hier vorgetrage-

nen Befunde nicht generalisiert werden. Es spricht unseres Erachtens aber wenig dafür, dass Bevölkerungsgruppen ohne Migrationshintergrund anders auf materielle Anreize reagieren. Wir betrachten hier allerdings junge Erwachsene und es erscheint daher durchaus möglich, dass die positive Wirkung des konditionalen Anteils der DI auch damit zu tun hat, dass es sich um eine Gruppe mit begrenzten finanziellen Mitteln handelt. Gegen dieses „Bedürftigkeitsargument“ spricht freilich der Befund, dass die Ausschöpfungsquote mit steigendem konditionalen Incentiveanteil nicht weiter zunimmt. Insgesamt betrachtet lassen unsere Befunde den Schluss zu, dass die doppelte Incentivierung als eine mögliche Alternative zur rein unkonditionalen Incentivierung künftig eine größere Aufmerksamkeit verdient.

Literatur

- AAPOR, 2011: Standard Definitions: The American Association for Public Opinion Research. Lenexa.
- Arzheimer, K. und M. Klein, 1998: Die Wirkung materieller Incentives auf den Rücklauf einer schriftlichen Panelbefragung. *ZA-Informationen* 43: 6-31.
- Aust, F. und H. Schröder, 2009: Sinkende Stichprobenausschöpfung in der Umfrageforschung – ein Bericht aus der Praxis. S. 195-212 in: M. Weichbold, J. Bacher und C. Wolf (Hg.), *Umfrageforschung. Herausforderungen und Grenzen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Babka von Gostomski, C. und M. Pupeter, 2008: Zufallsbefragung von Ausländern auf Basis des Ausländerzentralregisters. *Methoden - Daten - Analysen* 2: 149-177.
- Barón, J., R. Breunig, D. Cobb-Clark, T. Gørgens und A. Sartbayeva, 2009: Does the Effect of Incentive Payments on Survey Response Rates Differ by Income Support History? *Journal of Official Statistics* 25: 483-507.
- Baykara-Krumme, H., 2010: Die Berliner pairfam Studie. Ein Methodenbericht. Arbeitspapier 16 des DFG-Schwerpunktes "Beziehungs- und Familienentwicklungspanel".
- Becker, R., R. Imhof und G. Mehlkop, 2007: Die Wirkung monetärer Anreize auf den Rücklauf bei einer postalischen Befragung und die Antworten auf Fragen zur Delinquenz. *Methoden - Daten - Analysen* 1: 131-159.
- Berger, F., U. Grob, H. Fend und W. Lauterbach, 2005: Möglichkeiten zur Optimierung der Rücklaufquote in postalischen Befragungen. Bericht über die Vorstudie zum Forschungsprojekt Life. *Zeitschrift für Soziologie der Erziehung und Sozialisation* 25: 99-107.
- Blohm, M. und C. Diehl, 2001: Wenn Migranten Migranten befragen. Zum Teilnahmeverhalten von Einwanderern bei Bevölkerungsumfragen. *Zeitschrift für Soziologie* 30: 223-242.
- Blossfeld, H., K. Golsch und G. Rohwer, 2007: *Event History Analysis with Stata*. New York: Erlbaum.
- Bosnjak, M. und T. Tuten, 2003: Prepaid and Promised Incentives in Web Surveys. In: *Social Science Computer Review* 21: 208-217.
- Castiglioni, L., K. Pffor und U. Krieger, 2008: The Effects of Incentives on Response Rates and Panel Attrition. *Survey Research Methods* 2: 151-158.

- Church, A., 1993: Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis. *Public Opinion Quarterly* 57: 62-79.
- Diehl, C. und P. Fick, 2012: Deutschsein auf Probe. *Soziale Welt* 63: 339-360.
- Diekmann, A. und B. Jann, 2001: Anreizformen und Ausschöpfungsquoten bei postalischen Befragungen. Eine Prüfung der Reziprozitätshypothese. *ZUMA-Nachrichten* 48: 18-27.
- Dillman, D., 1978: *Mail and Telephone Surveys*. New York: Wiley.
- Dillman, D. 2007: *Mail and Internet Surveys*. Hoboken: Wiley.
- Dillman, D., J. Smyth und L. Christian, 2009: *Internet, Mail, and Mixed-Mode Surveys. The Tailored Design Method*, Hoboken: John Wiley & Sons.
- Dykema, J., J. Stevenson, B. Day, S. Sellers und V. Bonham, 2011: Effects of Incentives and Prenotification on Response Rates and Costs in a National Web Survey of Physicians. *Evaluation & the health Professions* 24: 434-447.
- Edwards, P., I. Roberts, M. Clarke, C. DiGiuseppi, S. Pratap, R. Wentz und I. Kwan, 2002: Increasing Response Rates to Postal Questionnaires. *British Medical Journal* 324: 1183-1185.
- Esser, H., 1986: Über die Teilnahme an Befragungen. *ZUMA-Nachrichten* 18: 38-47.
- Feskens, R., J. Hox, G. Lensvelt-Mulders und H. Schmeets, 2006: Collecting Data among Ethnic Minorities in an International Perspective. *Field Methods* 18: 284-304.
- Feskens, R., J. Hox, H. Schmeets und W. Wetzels, 2008: Incentives and Ethnic Minorities. *Survey Research Methods* 2: 159-165.
- Goebel, J., K. Spieß, N. Witte und S. Gerstenberg, 2007: Die Verknüpfung des SOEP mit MICROM-Indikatoren. *Data Documentation* 26. Berlin: DIW.
- Goto, S., 1996: To Trust or Not to Trust: Situational and Dispositional Determinants. *Social Behavior and Personality* 24: 119-131.
- Gouldner, A., 1960: The Norm of Reciprocity: A Preliminary Statement. *American Sociological Review* 25: 161-178.
- Goyder, J., 1994: An Experiment with Cash Incentives on a Personal Interview Survey. *Journal of the Market Research Society* 36: 360-366.
- Granato, N., 1999: Die Befragung von Arbeitsmigranten: Einwohnermeldeamt-Stichprobe und telefonische Erhebung? *ZUMA-Nachrichten* 45: 44-60.
- Grant, R. und J. Sugarman, 2004: Ethics in Human Subjects Research: Do Incentives Matter? *Journal of Medicine and Philosophy* 29: 717-138.
- Groves, R., E. Singer und A. Cornig, 2000: Leverage-Saliency Theory of Survey Participation. *Public Opinion Quarterly* 64: 299-308.
- Haberfeld, Y., Y. Cohen, F. Kalter und I. Kogan, 2011: Differences in Earnings Assimilation of Immigrants from the Former Soviet Union to Germany and Israel During 1994-2005. *International Journal of Comparative Sociology* 52: 6-24.
- Harkness, J., P. Mohler, M. Schneid und B. Christoph, 1998: Incentives in Two German Mail Surveys 1996/97 & 1997. S. 201-218 in: A. Koch und R. Porst (Hg.), *Nonresponse in Survey Research*. Mannheim: ZUMA.
- Haug, S. und C. Diehl, 2005: Aspekte der Integration. Wanderungsverhalten, Eingliederungsmuster und Lebenssituation türkisch- und italienischstämmiger Jugendlicher in Deutschland. *Schriftenreihe des Bundesinstituts für Bevölkerungsforschung*, Band 35.
- James, J. und R. Bolstein, 1992: Large Monetary Incentives and Their Effect on Mail Survey Response Rates. *Public Opinion Quarterly* 56: 442-453.
- James, J. und R. Bolstein, 1990: The Effect of Monetary Incentives and Follow-Up Mailings on the Response Rate and Response Quality in Mail Surveys. *Public Opinion Quarterly* 54: 346-361.
- Jobber, D., J. Saunders und V. Mitchell, 2004: Prepaid monetary incentive effects on mail survey response. *Journal of Business Research* 57: 21-25.

- Johnson, T., D. O'Rourke, J. Burris und L. Owens, 2002: Culture and Survey Nonresponse. S. 55-69 in: R. Groves, D. Dillman, J. Eltinge und R. Little (Hg.): Survey Nonresponse. New York: John Wiley & Sons.
- Lesser, V., D. Dillman, J. Carlson, F. Lorenz, R. Mason, F. Willits, 2001: Quantifying the Influence of Incentives on Mail Survey Response Rates and Their Effects on Nonresponse Error. Proceedings of the Annual Meeting of the American Statistical Association: 1-6.
- Mauss, M., 1968: Die Gabe. Frankfurt am Main: Suhrkamp.
- Mehlkop, G. und R. Becker, 2007: Zur Wirkung monetärer Anreize auf die Rücklaufquote in postalischen Befragungen zu kriminellen Handlungen. Methoden - Daten - Analysen 1: 5-24.
- Millar, M. und D. Dillman, 2011: Improving Response to Web and Mixed-Mode Surveys. Public Opinion Quarterly 75: 249-269.
- Paolillo, J. und P. Lorenzi, 1984: Monetary Incentives and Mail Questionnaire Response Rates. Journal of Advertising 13: 46-48.
- Porst, R., 1999: Thematik oder Incentives? Zur Erhöhung der Rücklaufquoten bei postalischen Befragungen. ZUMA-Nachrichten 45: 72-87.
- Porter, S. und M. Whitcomb, 2003: The Impact of Lottery Incentives on Student Survey Response Rates. Research in Higher Education 44: 389-407.
- Ryu, E., M. Couper und R. Marans, 2005: Survey Incentives: Cash vs. In-Kind; Face-to-Face vs. Mail; Response Rate vs. Nonresponse Error. International Journal of Public Opinion Research 18: 89-106.
- Schnell, R., 2012: Survey-Interviews. Methoden standardisierter Befragungen. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schupp, J., 2012: Die verborgenen Kosten monetärer Anreize – lohnt sich Motivierung durch Incentivierung? DIW Wochenbericht 79: 20.
- Schupp, J., N. Siegel, S. Huber und G. Wagner, 2010: Incentives and Response Rates. First Experiences from the SOEP-Innovation-Sample 2009. "2nd Panel-Survey-Methodology-Workshop", Mannheim (Vortragsfolien).
- Shettle, C. und G. Mooney, 1999: Monetary Incentives in U.S. Government Surveys. Journal of Official Statistics 15: 231-250.
- Singer, E., 1998: Incentives for Survey Participation: Research on Intended and Unintended Consequences. ZUMA-Nachrichten 42: 7-29.
- Singer, E., 2002: The Use of Incentives to Reduce Nonresponse in Household Surveys. S. 163-177 in: R. Groves, D. Dillman, J. Eltinge, R. Little (Hg.): Survey Nonresponse. New York: Wiley.
- Singer, E., J. Van Hoewyk, N. Gebler, T. Raghunathan und K. McGonagle, 1999: The Effects of Incentives on Response Rates in Interviewer-Mediated Surveys. Journal of Official Statistics 15: 217-230.
- Singer, E., J. Van Hoewyk und M. Maher, 2000: Experiments with Incentives in Telephone Surveys. Public Opinion Quarterly 64: 171-188.
- Stadtmüller, S., 2009: Rücklauf gut, alles gut? Zu erwünschten und unerwünschten Effekten monetärer Anreize bei postalischen Befragungen. Methoden - Daten - Analysen 3: 167-185.
- Stadtmüller, S. und R. Porst, 2005: Zum Einsatz von Incentives bei postalischen Befragungen. ZUMA How-to-Reihe 14.
- Statistisches Bundesamt, 2011: Fachserie 1 Reihe 2.2, Wiesbaden: StBA.
- Trussell, N. und P. Lavrakas, 2004: The Influence of Incremental Increases in Token Cash Incentives on Mail Survey Response. Public Opinion Quarterly 68: 349-367.

- Warriner, K., J. Goyder, H. Gjertsen, P. Hohner und K. McSpurren, 1996: Charities, No; Lotteries, No; Cash, Yes. Main Effects and Interactions in a Canadian Incentive Experiment. *Public Opinion Quarterly* 60: 542-562.
- Weinmann, M. I. Becher und C. Babka von Gostomski, 2012: Einbürgerungsverhalten von Ausländerinnen und Ausländern in Deutschland sowie Erkenntnisse zu Optionspflichtigen. Nürnberg: BAMF.
- Wetzels, W., H. Schmeets, J. van den Brakel und R. Feskens, 2008: Impact of Prepaid Incentives in Face-to-Face Surveys: A Large-Scale Experiment with Postage Stamps. *International Journal of Public Opinion Research* 20: 507-516.
- Wotruba, T., 1966: Monetary Inducements and Mail Questionnaire Response. *Journal of Marketing Research* 3: 398-400.
- Yu, J. und H. Cooper, 1983: A Quantitative Review of Research Design Effects on Response Rates to Questionnaires. *Journal of Marketing Research* 20: 36-44.
- Zentrum für Türkeistudien, 2012: Integrationsprozesse türkeistämmiger Migrantinnen und Migranten in Nordrhein-Westfalen. Essen: ZfT.

Anschrift der Autoren

Patrick Fick
Universität Konstanz
Fachbereich Geschichte und Soziologie
Universitätsstr. 10
78457 Konstanz
E-Mail: patrick.fick@uni-konstanz.de

Claudia Diehl
Universität Konstanz
Fachbereich Geschichte und Soziologie
Universitätsstr. 10
78457 Konstanz
E-Mail: claudia.diehl@uni-konstanz.de