

Zusatzkapitel 4: Multiple Imputation des wahrgenommenen Risikos in der Analyse der Rettung von Juden im Zweiten Weltkrieg^{*}

Für „unbeteiligte“ Befragte, also diejenigen Personen, die weder mit einem Hilfesuch konfrontiert, noch eigeninitiativ wurden, liegt keine Messung des generell wahrgenommenen Risikos oder des wahrgenommenen Risikos für die eigene Familie vor. Im Folgenden beschreibe ich die multiplen Imputationsverfahren, mit deren Hilfe diesen Fällen dennoch Werte(-verteilungen) auf diesen Variablen zugewiesen wurden. Das Ziel bestand darin, die Fallzahl und damit die Teststärke der Analysen möglichst hoch zu halten und gleichzeitig eine möglichst vollständige Analyse der Determinanten der Hilfeleistung zu ermöglichen. Das eingesetzte Verfahren wird als „multiple imputation by chained equations“ (kurz: MICE) bezeichnet (Royston 2005a, b; Van Buuren, Boshuizen und Knook 1999).

Bei multipler Imputation wird zunächst ein Regressionsmodell spezifiziert, mit dessen Hilfe Werte auf der zu imputierenden Variablen vorhergesagt werden können. In dieses sog. *Imputationsmodell* können selbstverständlich nur Fälle mit beobachteten Werten auf der zu imputierenden Variablen eingehen. Es kann aber – unter bestimmten Annahmen – dazu verwendet werden, Werte für diejenigen Fälle vorherzusagen, bei denen entsprechende Werte fehlen.

Im Gegensatz zu primitiveren Verfahren des Umgangs mit fehlenden Werten werden diese Imputationen nicht genauso wie gemessene Werte behandelt. Vielmehr wird berücksichtigt, dass die Vorhersagen auf Basis des Imputationsmodells mit Unsicherheit behaftet sind – wobei das Ausmaß der Unsicherheit von der Güte des Imputationsmodells abhängt. Dies wird

^{*} Dieses Zusatzkapitel kann wie folgt zitiert werden: Kroneberg, Clemens (2011): Zusatzkapitel 4 zu „Die Erklärung sozialen Handelns“: Multiple Imputation des wahrgenommenen Risikos in der Analyse der Rettung von Juden im Zweiten Weltkrieg. URL: <http://vs-verlag.de/tu/Kroneberg-Erklaerung>. Stand: 01.06.2011

erreicht, indem auf Basis der Fehlerstreuung des Imputationsmodells für jeden Fall eine *Vielzahl* von Imputationen vorgenommen wird. So entsteht eine Vielzahl vervollständigter Datensätze, in denen die fehlenden Werte durch imputierte ersetzt sind. Im vorliegenden Anwendungsfall wurden fünfzig derart vervollständigte Datensätze erstellt, um den Einfluss zufälliger Schwankungen zu minimieren.

Für jeden dieser vervollständigten Datensätze wird nun das letztlich interessierende Regressionsmodell geschätzt. Man erhält so über die Vielzahl der Datensätze hinweg eine *Verteilung* geschätzter Regressionskoeffizienten und Standardfehler. Aus diesen Verteilungen lassen sich mit Hilfe von Rubins Regeln für jeden Prädiktor jeweils ein mittlerer Regressionskoeffizient und ein mittlerer Standardfehler berechnen: Der resultierende Regressionskoeffizient entspricht dem arithmetischen Mittel seiner Verteilung über alle vervollständigten Datensätze. In seinen Standardfehler geht sowohl die Variation innerhalb jeder Imputation (gewöhnliche Schätzunsicherheit) als auch die zwischen den Imputationen (Unsicherheit durch Imputation) ein.

Multiple Imputation basiert auf der Annahme, dass das Fehlen der Werte aus einem Zufallsprozess resultiert (MAR = „missing at random“). Von zufällig fehlenden Werten kann im vorliegenden Anwendungsfall aber nicht ausgegangen werden, denn das wahrgenommene Risiko fehlt systematisch für all diejenigen Fälle, bei denen kein Hilfesuch vorlag und in denen es keine Eigeninitiative gab. Es kann daher nicht ohne Weiteres angenommen werden, dass die Risikowahrnehmung dieser Personen in gleicher Weise zustande kam wie die von Personen, die um Hilfe gebeten wurden oder eigeninitiiert halfen.

Bei der multiplen Imputation des wahrgenommenen Risikos auf Basis der Fälle mit gemessenen Werten muss daher zusätzlich die Selektivität dieser Teilstichprobe statistisch berücksichtigt werden. Dies geschieht mit Hilfe des im Haupttext berichteten multinomialen logistischen Regressionsmodells, das vorhersagt, ob es zu einem Hilfesuch, eigeninitiiert oder weder dem einen noch dem anderen („unbeteiligt“) kam. Vor die Imputationsmodelle des wahrgenommenen Risikos wird somit gleichsam ein Selektivitätsmodell geschaltet. Die Wahrscheinlichkeit, mit einem Hilfesuch konfrontiert worden zu sein oder eigeninitiativ geholfen zu haben, lässt sich dann innerhalb der Imputationsmodelle für das wahrgenommene Risiko als zusätzliche unabhängige Variable berücksichtigen. Denn obwohl alle in das Imputationsmodell eingehenden Fälle entweder angesprochen oder eigenini-

tativ wurden, hatten sie doch eine systematisch unterschiedliche hohe Wahrscheinlichkeit dafür.

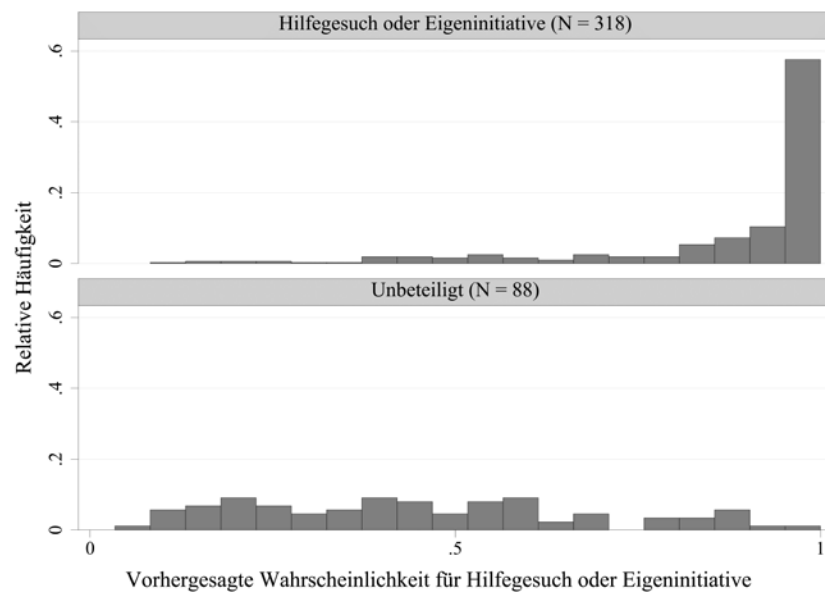
Dies wird in Abbildung 1 deutlich, welche die vorhergesagte Wahrscheinlichkeit, mit einem Hilfesuch konfrontiert worden zu sein oder eigeninitiativ geholfen zu haben, zwischen den Fällen, in denen das eine oder andere eintrat, und den „unbeteiligten“ Fällen vergleicht. Einerseits unterscheiden sich die Verteilungen deutlich. Die vorhergesagte Wahrscheinlichkeit liegt bei Fällen mit Hilfesuch oder eigeninitiiertter Hilfe realistischerweise durchschnittlich deutlich höher. Andererseits überschneiden sich die Verteilungen hinreichend, um die angestrebte multiple Imputation vornehmen zu können. Das heißt, es existieren genügend Fälle mit Hilfesuch oder eigeninitiiertter Hilfe, bei denen dies eigentlich relativ unwahrscheinlich war, sowie genügend „unbeteiligte“ Fälle, bei denen eine Ansprache oder eigeninitiiertes Helfen durchaus wahrscheinlich gewesen wäre.

Indem der Einfluss dieser vorhergesagten Wahrscheinlichkeiten in den Imputationsmodellen statistisch herauspartialisiert wird, soll im Rahmen der multiplen Imputation der Risikoindikatoren für die Selektivität der Teilstichprobe mit Risikomessung im Vergleich zur Gesamtstichprobe kontrolliert werden.

Wie Heckman gezeigt hat, sind Verzerrungen durch Selektivität äquivalent zu Fehlspezifikationen durch Nicht-Berücksichtigung eines relevanten Prädiktors (Heckman 1979). Durch Einschluss einer geeigneten *Transformation* der vorhergesagten Wahrscheinlichkeit kann die andernfalls bestehende Korrelation zwischen dem Fehlerterm und den anderen Prädiktoren prinzipiell wieder orthogonalisiert werden. Für den Fall, dass das substantiell interessierende Modell eine lineare Regression ist, hat Heckman die geeignete Transformation der vorhergesagten Wahrscheinlichkeit hergeleitet (die sog. inverse Mills ratio). Für den hier vorliegenden Anwendungsfall, in dem zuerst eine multinomiale logistische Regression als Selektivitätsmodell und danach eine ordinale logistische Regression als substantiell interessierendes Modell geschätzt werden müssen, existiert jedoch bislang keine theoretisch abgesicherte Vorgehensweise.¹

¹ Die jüngst entwickelten Verfahren (Greene 2006; Miranda und Rabe-Hesketh 2006) sind hier nicht anwendbar, da sie auf der Maximierung einer gemeinsamen Likelihood-Funktion beruhen. Da das in unserem Kontext substantiell interessierende Regressionsmodell (des wahrgenommenen Risikos) ein multiples Imputationsmodell darstellt, ist ein

Abbildung 1: Verteilung der vorhergesagten Wahrscheinlichkeit, mit einem Hilfesuch konfrontiert oder eigeninitiativ zu werden, für unbeteiligte Fälle und Fälle mit Hilfesuch oder Eigeninitiative



Analog zum Vorgehen in anderen Studien (z.B. Morgan und Sorensen 1999) wurde daher getestet, ob die Selektionskorrektur eine nicht-lineare Transformation der vorhergesagten Wahrscheinlichkeit erfordert. Diese Robustheitsanalysen ergaben jedoch keine statistisch signifikanten Hinweise auf einen derartigen nicht-linearen Einfluss.² Zudem erwiesen sich die Ergeb-

sog. two-step-Verfahren, in dem zuerst das Selektivitätsmodell separat geschätzt wird, unumgänglich.

² Erstens wurde auf Basis des multinomialen Selektivitätsmodells die inverse Mills ratio berechnet und als Prädiktor in den Imputationsmodellen verwendet. Theoretisch ist dies jedoch (unter der Annahme bivariat normalverteilter Fehlerterme) nur gerechtfertigt, wenn das substantiell interessierende Modell eine lineare Regression ist. Eine Übertragung auf andere Modelle ist zwar weit verbreitet, aber nicht begründet (Bushway, Johnson und Slocum 2007). Zweitens wurde daher eine nicht-lineare Transformation der vor-

nisse der letztendlich interessierenden Hypothesentests als robust gegenüber derartigen Variationen. Der einfacheren Interpretierbarkeit halber werden daher lediglich die Resultate unter Einschluss der (nicht-transformierten) vorhergesagten Wahrscheinlichkeit berichtet.

Die Spezifikation des Selektivitätsmodells und der Imputationsmodelle erfolgte auf Basis von Plausibilitätserwägungen und einer Vorinspektion potentiell einflussreicher Prädiktoren (vgl. Van Buuren, Boshuizen und Knook 1999: 687). Im Rahmen letzterer wurden Prädiktoren ohne jegliche Erklärungskraft (Regressionskoeffizient nahe null) und mit extrem hohem Anteil fehlender Werte ausgeschlossen. Eine Optimierung der Regressionsmodelle unter Einbezug des Kriteriums statistischer Signifikanz (sog. „model trimming“) wurde dagegen bewusst *vermieden*, da dies bekanntlich zu einer Überanpassung an die Stichprobe und darüber zu Verzerrungen führt. Dies gilt gerade auch im Kontext von Selektivitätsmodellen, mit deren Hilfe die Wahrscheinlichkeit einer Selektion vorhergesagt werden soll (Rubin und Thomas 1996).

Alle verwendeten Variablen, die nicht bereits in die Hauptanalysen eingingen, sind in Tabelle 2 am Ende dieses methodischen Anhangs aufgeführt. Mit wenigen Ausnahmen gehen alle im Selektivitätsmodell berücksichtigten Prädiktoren auch in die Imputationsmodelle des wahrgenommenen Risikos ein. Die Ausnahmen dienen als sog. *Exklusionsrestriktionen*, denen bei der statistischen Modellierung von Selektivität eine zentrale Rolle zukommt. Dies sind Variablen, von denen angenommen werden kann, dass sie zwar die Selektion beeinflussen, nicht aber die substantiell interessierende abhängige Variable (Bushway, Johnson und Slocum 2007: 163ff.). In unserem Anwendungskontext geht es also um Variablen, die beeinflussten, ob man unbeteiligt blieb (im Unterschied zur Konfrontation mit einem Hilfesuch oder eigeninitiiertem Hilfe), die aber *nicht* mit systematisch unterschiedlichem wahrgenommenen Risiko einhergingen. Beispielsweise kommen Indikatoren objektiv günstiger Gelegenheiten zur Hilfe *nicht* in Frage, da sie nicht nur Hilfesuche oder Eigeninitiative wahrscheinlicher gemacht haben dürften,

hergesagten Wahrscheinlichkeit eines Hilfesuchs in Form von bis zu drei Prädiktoren berücksichtigt. Diese Transformation wurde mit Hilfe des sehr flexiblen *fracpoly*-Verfahrens (siehe StataCorp. 2007) so gewählt, dass sie am besten auf die Daten passt. Allerdings ergab auch dieses Verfahren einen weitgehend linearen Einfluss und entsprechend keine statistisch signifikante Verbesserung gegenüber dem Einschluss der nicht-transformierten Wahrscheinlichkeit.

sondern theoretisch bei besseren Gelegenheiten auch mit einer durchschnittlich niedrigeren subjektiven Risikowahrnehmung zu rechnen ist. Geeignet sind dagegen Variablen, die ausschließlich mit der Präferenz oder generellen Disposition zur Hilfeleistung zusammenhängen.

Auf Basis dieser Überlegung erscheint es theoretisch gerechtfertigt, für die folgenden drei Variablen von der gesuchten Beziehungsstruktur auszugehen: Erstens dürfte die generelle *Hilfsbereitschaft* in der Kindheit damit korrelieren, ob man eigeninitiativ oder angesprochen wurde und wie man sich im Falle eines Hilfesuchs verhielt. Indes besteht kein Grund anzunehmen, dass hilfsbereitere Personen die Situation generell als riskanter oder weniger riskant einschätzten. Gleiches gilt zweitens für die Selbsteinschätzung des Befragten, inwieweit er bereits in seiner Kindheit fähig war, Verantwortung zu übernehmen (*Verantwortungsvoll*). Auch diese Disposition steht theoretisch in keinem direkten Bezug zur Risikowahrnehmung, wohl aber zur Wahrscheinlichkeit, eigeninitiiert zu helfen oder angesprochen zu werden.³ Drittens wurden die Antworten auf die Frage verwendet, ob der Befragte bereits vor Kriegsbeginn außergewöhnliches Hilfeverhalten gezeigt habe (*Früheres Hilfeverhalten*). Die Risikowahrnehmung könnte höchstens durch die spezifischen Erfahrungen bei diesem früheren Hilfeverhalten geprägt sein. Da diese Erfahrungen jedoch stark variiert haben dürften, ergibt sich theoretisch wiederum keine eindeutige Einflussrichtung auf das wahrgenommene Risiko.

Mit Ausnahme dieser drei Exklusionskriterien gehen alle im Selektivitätsmodell verwendeten Prädiktoren auch in die Imputationsmodelle des wahrgenommenen Risikos ein. Umgekehrt wird jedoch eine Reihe von Variablen ausschließlich als Prädiktoren des Risikos verwendet, da sie sich im Selektivitätsmodell sowohl substantiell als auch bivariat als vernachlässigbar erwiesen. Dies sind das Alter im Quadrat (zur Modellierung eines nicht-linearen Einflusses auf die Risikowahrnehmung), Merkmale des familiären Kontextes während der Kindheit (Alleinerziehender Vater, Geschwisteranzahl, Jüngere Geschwister), die Stellung im Haushalt während des Krieges (Hauptverdiener), der Wohnort (Stadt), die Identifikation mit Juden und das Vorhandensein jüdischer Nachbarn. Wie bereits erläutert, wurde zum Zwe-

³ Auf Basis des gleichen Arguments könnte auch die prosoziale Orientierung als Exklusionskriterium verwendet werden. Dies ist allerdings nicht möglich, da dieser Prädiktor aus anderen Gründen auch bei der Imputation des wahrgenommenen Risikos berücksichtigt werden muss (siehe dazu noch unten).

cke der Selektivitätskorrektur zudem die vorhergesagte Wahrscheinlichkeit, nicht unbeteiligt gewesen zu sein (Prob(„nicht unbeteiligt“)), mit aufgenommen. Auch wurde für das tatsächliche Vorliegen eines Hilfesuchs kontrolliert, um die diesbezügliche Heterogenität der Teilstichprobe mit Risikomessung zu berücksichtigen. Hervorzuheben sind schließlich zwei weitere Prädiktoren, denen eine besondere statistische Bedeutung im Rahmen der multiplen Imputation zukommt: Das eigene Hilfeverhalten sowie ein Produktterm zwischen dem eigenen Hilfeverhalten und der prosozialen Orientierung.⁴

Dass in die Imputationsmodelle des wahrgenommenen Risikos als Prädiktor auch das Hilfeverhalten selbst eingeht, mag auf den ersten Blick zirkulär erscheinen. Denn im letztendlich substantiell interessierenden Regressionsmodell geht es ja darum, das Hilfeverhalten *umgekehrt* durch das wahrgenommene Risiko zu erklären. Tatsächlich besteht aber keine Zirkularität, da keine bedingten Mittelwerte imputiert werden, sondern (mehrfach) Zufallswerte auf Basis der bedingten (posterior-)Verteilung gezogen werden (Little und Rubin 2002: 66). Zu Verzerrungen käme es im Gegenteil dann, wenn die letztlich interessierende abhängige Variable *nicht* als Prädiktor in die Imputationsmodelle aufgenommen würde (Little 1992: 1235).

Ebenso ist bei der Imputation zu berücksichtigen, dass theoretisch Grund zu der Annahme besteht, dass die prosoziale Orientierung die Beziehung zwischen dem wahrgenommenen Risiko und dem Hilfeverhalten moderiert. Das macht es notwendig, bei der Imputation des wahrgenommenen Risikos einen analogen Interaktionseffekt zwischen dem Hilfeverhalten und der prosozialen Orientierung zu modellieren.⁵ Bei der Imputation fehlender Werte auf anderen Anreizvariablen wurde daher zusätzlich auch noch ein Produktterm zwischen dem Vorliegen eines Hilfesuchs und dem Hilfeverhalten berücksichtigt.

⁴ Das Vorliegen eines Hilfesuchs sowie teilweise das eigene Hilfeverhalten gingen in das Selektivitätsmodell der Ansprache allerdings bereits in Form der kategorialen abhängigen Variablen ein.

⁵ Dadurch wird zugelassen, dass die Beziehung zwischen Hilfeverhalten und wahrgenommenem Risiko je nach Ausprägung der prosozialen Orientierung variiert. Andernfalls würden möglicherweise bestehende Interaktionseffekte systematisch unterschätzt. Es läge also eine Fehlspezifikation des multiplen Imputationsmodells vor. Dies könnte das Ergebnis einer aktuellen Monte-Carlo-Studie zur hier verwendeten Methode erklären, der zufolge Nicht-Linearitäten von allen getesteten Imputationsverfahren verdeckt wurden (Hardt und Görgen 2008).

Für das generelle wahrgenommene Risiko und das Risiko für die eigene Familie wurden getrennte Imputationsmodelle spezifiziert. Es ist davon auszugehen, dass diese Indikatoren teilweise unterschiedliche Korrelate aufweisen. Zudem sollte identisch vorgegangen werden wie bei Fällen mit gemessenen Werten. Auch die imputierten Werte auf den beiden Risiko-Indikatoren wurden daher erst im Nachhinein zu einem ungewichteten additiven Index zusammengefasst. Dies geschah wiederum nur bei Fällen mit Familienangehörigen im Haushalt. Ansonsten wurde nur das generell wahrgenommene Risiko verwendet. Bei Fällen mit fehlenden Werten auf nur einem der beiden Risikoindikatoren wurde nur die vorhandene Risikomesung verwendet.

Tabelle 1 enthält die Ergebnisse der Imputationsmodelle für die beiden Risikoindikatoren. Da wiederum bewusst keine Modellselektion nach statistischen Signifikanzkriterien vorgenommen wurde, ist die überwiegende Zahl der Koeffizienten statistisch insignifikant. Konsistente, statistisch signifikante Effekte auf *beide* Risikoindikatoren gehen von drei Prädiktoren aus: Je mehr finanzielle Ressourcen dem Haushalt des Befragten zur Verfügung standen, umso geringer war das wahrgenommene Risiko. Die Existenz vieler Nachbarn erhöhte dagegen das wahrgenommene Risiko. Eine systematisch niedrigere Risikowahrnehmung hatten interessanterweise Befragte, die nur mit ihrem Vater aufwuchsen. Daneben ergeben sich eine Reihe einzelner signifikanter Zusammenhänge für das generelle bzw. das für die eigene Familie wahrgenommene Risiko.

Das generelle wahrgenommene Risiko war tendenziell höher bei ausgeprägter Identifikation mit den Juden und sofern Kinder im Haushalt lebten. Es war dagegen durchschnittlich niedriger bei Befragten mit jüdischen Nachbarn und solchen, deren Familienangehörige im Widerstand aktiv waren oder Juden halfen, sowie tendenziell, je mehr Räume zur Verfügung standen. Bei Fällen, in denen es zur Hilfeleistung kam, war das wahrgenommene Risiko durchschnittlich geringer, sofern eine geringe prosoziale Orientierung vorlag.

Das wahrgenommene Risiko für die eigene Familie war höher bei Befragten, die sich in ihrer Kindheit für ihre Überzeugungen eingesetzt hatten (frühere Überzeugungstaten) oder sich politisch engagiert hatten. Zudem war es durchschnittlich höher bei einer umfassenderen Kontrolle des Landes durch die Nazis. Es war dagegen durchschnittlich niedriger bei Befrag-

ten, die Hauptverdiener ihres Haushalts waren, und solchen, die Misshandlungen durch die Nazis wahrgenommen hatten.

Von der vorhergesagten Wahrscheinlichkeit eines Hilfesuchs oder eigeninitiiertem Helfen geht in beiden Modellen kein statistisch signifikanter Einfluss aus. Daraus sollte allerdings nicht geschlossen werden, dass die Selektivitätskorrektur unnötig oder nicht erfolgreich war. Bei Standardanwendungen statistischer Selektivitätskorrektur bleibt die Stichprobe, in der das letztlich interessierende Modell geschätzt wird, selektiv und es wird durch einen entsprechenden Term (im klassischen Anwendungsfall: die sog. inverse Mills ratio) für diese Selektivität statistisch kontrolliert. Im hier vorliegenden Anwendungskontext trifft diese Beschreibung dagegen lediglich *im Prozess* der multiplen Imputation der Risikoindikatoren zu: Die schrittweise und wiederholte Vorhersage von Risikowerten berücksichtigt die Selektivität der Teilstichprobe, auf der diese Vorhersage beruht. Ist diese multiple Imputation dagegen *abgeschlossen* – wie in den in Tabelle 1 dargestellten Modellen –, so ist von keiner Selektivität mehr auszugehen, da nun auch diejenigen Fälle in das Modell eingehen, für die keine Risikomessung vorliegt.

Tabelle 1: Prädiktoren des generellen und für die eigene Familie wahrgenommenen Risikos (Ordinale logistische Regressionen)

	Risiko generell		Risiko Familie	
Prob(„nicht unbeteiligt“)	1.32	(1.56)	-0.68	(1.40)
Hilfesuch	-0.16	(0.39)	0.11	(0.41)
Hilfeverhalten	-3.31*	(1.52)	-0.78	(1.54)
Prosoz. Orientierung (PO)	-0.93	(2.36)	-0.41	(2.44)
Hilfeverhalten × PO	1.74	(2.34)	-0.35	(2.28)
Frau	0.09	(0.36)	-0.02	(0.46)
Alter (in Jahren)	-0.03	(0.14)	-0.20	(0.15)
Alter quadriert	0.00	(0.00)	0.00	(0.00)
abgeschl. Studium	0.19	(0.48)	-0.57	(0.54)
abgeschl. Gymnasium	0.38	(0.52)	0.05	(0.54)
abgeschl. Lehre	-0.50	(0.54)	0.15	(0.50)
Partner im HH	0.07	(0.60)	0.48	(0.57)
Kinder im HH	0.79+	(0.47)	0.12	(0.51)

Fortsetzung von vorhergehender Seite

Hauptverdiener	-0.22	(0.35)	-0.74+	(0.40)
Frühere Überzeugungstaten	0.44	(0.38)	1.36**	(0.47)
Früheres polit. Engagement	0.34	(0.46)	0.78+	(0.42)
Früheres Risikoverhalten	-0.08	(0.44)	0.46	(0.49)
Selbstvertrauen	0.84	(0.59)	-0.24	(0.67)
Nur Vater	-3.43**	(1.07)	-3.86**	(1.03)
Anzahl Geschwister	-0.05	(0.08)	-0.08	(0.08)
Jüngere Geschwister	0.09	(0.39)	0.51	(0.41)
#Räume	-1.01	(0.69)	0.05	(0.73)
Finanzielle Ressourcen	-3.30**	(0.93)	-1.75+	(0.93)
Viele Nachbarn	0.94*	(0.44)	0.89+	(0.50)
Stadt	0.10	(0.41)	-0.44	(0.45)
Kontakt zu Juden	-0.21	(0.41)	0.10	(0.38)
Misshandlungen:				
Wahrgenommen	-0.20	(0.43)	-0.85+	(0.50)
Erlitten	0.50	(0.56)	0.15	(0.52)
Identifikation mit Juden	1.17*	(0.46)	0.44	(0.38)
Jüdische Nachbarn	-0.86*	(0.40)	-0.11	(0.39)
Widerstand	0.60	(0.46)	0.06	(0.45)
Widerstand/Hilfe in Familie	-0.88*	(0.43)	-0.05	(0.36)
Befehlszone	0.10	(0.71)	2.53**	(0.74)
SS-Zone	-1.50+	(0.80)	3.22**	(0.81)
%Jüdische Bevölkerung	0.16	(0.25)	0.28	(0.32)
SS-Zone × %JB	1.10	(2.37)	-2.08	(2.92)
τ_1	-5.05+	(2.86)	-4.11	(2.94)
τ_2	-4.27	(2.84)	-3.67	(2.94)
τ_3	-2.96	(2.83)	-2.54	(2.91)
N	406		406	

Anmerkungen: Logit-Koeffizienten; in Klammern: robuste Standardfehler mit Berücksichtigung der Cluster

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ (zweiseitige Tests)

Neben dem konsistenten und starken Einfluss der Existenz vieler Nachbarn, der finanziellen Ressourcen und des Aufwachsens bei einem alleinerziehenden Vater ergibt sich somit ein relativ disparates Bild unterschiedlicher und nicht immer leicht zu interpretierender Zusammenhänge. Allerdings geht es hier auch nicht um eine theoriegeleitete Erklärung der Risikowahrnehmung, sondern um eine multivariate Abbildung der in den Daten bestehenden Zusammenhänge zum Zwecke der Imputation fehlender Werte.

In Abbildung 2 ist für beide Risiko-Indikatoren die Verteilung der beobachteten Werte der Verteilung der imputierten Werte (über alle fünfzig Datensätze hinweg) gegenübergestellt. An der weitgehenden Übereinstimmung der Verteilungen wird deutlich, dass die multiple Imputation fehlende Werte durch realistische Werteverteilungen ersetzt (im Gegensatz zu primitiveren Imputationsverfahren).

Abbildung 2: Verteilungen der gemessenen und imputierten Werte auf den beiden Indikatoren des wahrgenommenen Risikos

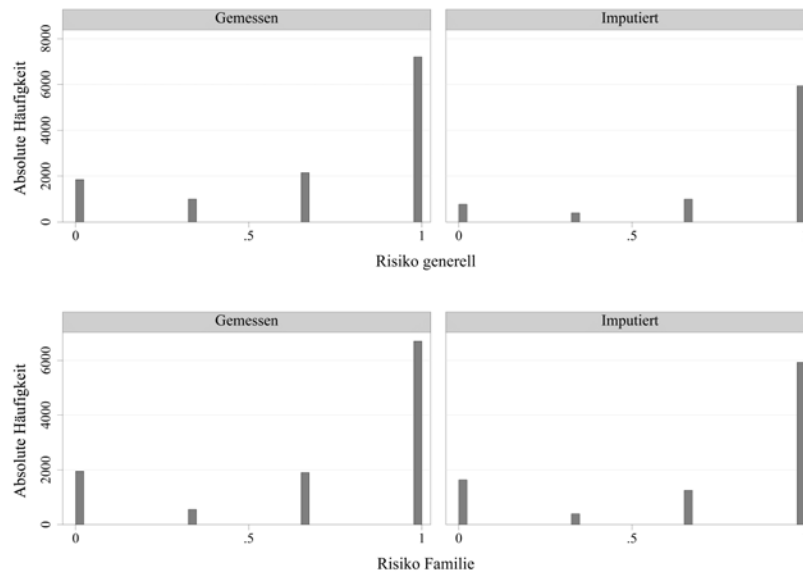


Tabelle 2: Überblick über die zusätzlichen, im Rahmen der Imputation des wahrgenommenen Risikos verwendeten Variablen
(Analysestichprobe mit N = 406)

Label	Frage ¹	Kodierung	Mittelwert (Std.abw.)	Fehlende Werte: Absolut (%)
Persönliche Eigenschaften:	„Now I would like to name some personal qualities. For each quality listed on page 7, I'd like to know if you thought that you had the quality very much, somewhat, not very much, or not at all <i>when you were growing up</i> .“ C18	1 = „very much“ 0.67 = „some-what“ 0.33 = „not very much“ 0 = „not at all“		
<i>Selbstvertrauen</i>	- “Self-confidence“ C18b		.67 (.31)	15 (3.69)
<i>Verantwortungsvoll</i>	- “Able to take responsibility“ C18d		.84 (.25)	5 (1.23)
<i>Hilfsbereitschaft</i>	- “Helpful to others“ C18h		.91 (.17)	10 (2.46)
<i>Kontakt zu Juden</i>	„Did you [your husband/wife] have any Jewish colleagues or coworkers in your [his/her] job?“ D11b [D3b] „Did you [your husband/wife] have any Jewish friends before the war?“ D12 [D4]	1 = mindestens einer dieser Kontakte besteht 0 = kein Kontakt	.59 (.49)	0 (0.00)
<i>Frühere Überzeugungen</i>	„Now I would like you to think about things that you may have done that were out of the ordinary, either while you were growing up or before the war. First of all, at any time in your life before the war, did you ever do anything unusual to stand up for your beliefs? For example, did you ever fight someone for something you believed in, or did you lose friends for standing up for what you believed in?“ D17	1 = „yes“ 0 = „no“	.32	18 (4.43)
<i>Früheres Hilfeverhalten</i>	„At any time in your life before the war, did you ever do anything unusual to help other people? For example, did you work as a volunteer in your community or in a hospital, or did you take a child into your home?“ D18	1 = „yes“ 0 = „no“	.30	7 (1.72)

Fortsetzung von vorhergehender Seite

Früheres polit. Engagement	„At any time in your life before the war, did you ever do anything unusual to defend or support your country? For example, were you active in any kind of political youth or student group?“ D19	1 = „yes“ 0 = „no“	.26	10 (2.46)
Früheres Risikoverhalten	„At any time in your life before the war, did you ever do anything unusual that was risky or adventurous?“ D20	1 = „yes“ 0 = „no“	.23	12 (2.96)
Identifikation mit Juden	„People sometimes think of others as being like or unlike themselves. [...] I would like you to think back to the time before the war. For each group of people, I'd like you to tell me what you thought of those people before the war. In general, did you think the people were very much like you, somewhat like you, not very much like you, or not like you at all?“ D21 „Jews“ D21f	1 = „Jews very much like me“ 0 = Ansonsten	.36	53 (13.05)
Ausmaß deutscher Kontrolle: SS-Zone Befehlszone	„Where were you living when the war broke out in your country? Which town and which country?“ E3	SS-Zone (= 1) Befehlszone (= 1) Ref.Kat.: Kolonialzone	.52 .33	0 (0.00)
%jüdische Bevölkerung	„Where were you living when the war broke out in your country? Which town and which country?“ E3	Anteil der jüdischen Bevölkerung in Prozent	3.78 (4.11)	0 (0.00)
Erlittene Misshandlungen	„Were there any times, other than those you may have already told me about, when you, personally, were mistreated by the Nazis?“ E56	1 = „yes“ 0 = „no“	.20	14 (3.45)
Wahrgenommene Misshandlungen	„Were there any times, other than those you may have already told me about, when you witnessed or heard about mistreatment of Jews by the Nazis?“ E57	1 = „yes“ 0 = „no“	.63	12 (2.96)
Widerstand/Hilfe in Familie	„Did your father or mother or any of your brothers or sisters engage in any kind of helping or resistance activities during the war?“ E59	1 = „yes“ 0 = „no“	.48	8 (1.97)
Haushaltstyp: Partner in HH	„Now I would like to make a list of the people who were living in your house when the war broke out in your country.“ E69a; Angabe der Beziehung aller Personen zum Befragten E69b	Partner im HH (= 1) Kinder im HH (= 1) Ref.Kat.: weder Partner noch Kinder im HH	.15 .37	58 (14.29)

Fortsetzung von vorhergehender Seite

Hauptverdiener	„Who was the head of your household? That is, who was the main breadwinner or wage earner?“ E69c	1 = Befragter 0 = Ansonsten	.36	0 (0.00)
Nur Vater	„Now, I would like you to think of your home when you were 10 years old. [...] Were your father and mother both living in your home when you were 10?“ A6	1 = Nur mit Vater 0 = Ansonsten	.03	0 (0.00)
Anzahl Geschwister	Summe der Anzahl älterer Brüder (A7a), älterer Schwestern (A7b), jüngerer Brüder (A8a) und jüngerer Schwestern (A8b)	15 Ausprägungen (0-1-kodiert): 0,1,2,3,...,13, 14	2.97 (2.57)	4 (0.99)
Jüngere Geschwister	„Did you have any younger brothers or sisters living with you when you were 10?“ A8	1 = „yes“ 0 = „no“	.63	4 (0.99)
Prob(„nicht unbeteiligt“)	Auf Basis des multinomialen logistischen Regressionsmodells (mit Imputation; siehe Tabelle 20 im Buch) vorhergesagte Wahrscheinlichkeit, nicht unbeteiligt gewesen zu sein, d.h., angesprochen worden zu sein oder eigeninitiiert geholfen zu haben.		.78 (.27)	0 (0.00)

Anmerkungen: ¹Die Identifikationsnummern der Fragen markieren ihre Position im Fragebogen (durchgängige Nummerierung innerhalb der Abschnitte A-E) (siehe Oliner und Oliner 1988: Appendix C)

Literatur

- Bushway, Shawn, Brian D. Johnson und Lee Ann Slocum, 2007: Is the Magic Still There? The Use of the Heckman Two-Step Correction for Selection Bias in Criminology. *Journal of Quantitative Criminology* 23: 151–178.
- Greene, William, 2006: "A General Approach to Incorporating Selectivity in a Model." in *Working Papers 06-10*. New York: New York University, Leonard N. Stern School of Business, Department of Economics.
- Hardt, Jochen und Kai Görden, 2008: "Multiple imputation using ICE: A simulation study on a binary response." in *6th German Stata Users Group meeting*. Berlin.
- Heckman, James J., 1979: Sample selection bias as a specification error. *Econometrica* 47: 153-161.
- Little, Roderick J. A., 1992: Regression With Missing X's: A Review. *Journal of the American Statistical Association* 87: 1227-1237.
- Little, Roderick J. A. und Donald B. Rubin, 2002: *Statistical Analysis with Missing Data*. New York: Wiley.
- Miranda, Alfonso und Sophia Rabe-Hesketh, 2006: Maximum likelihood estimation of endogenous switching and sample selection models for binary, ordinal, and count variables. *The Stata Journal* 6: 285-308.

- Morgan, Stephen L. und Aage B. Sorensen, 1999: Parental Networks, Social Closure, and Mathematics Learning: A Test of Coleman's Social Capital Explanation of School Effects. *American Sociological Review* 64: 661-681.
- Oliner, Samuel P. und Pearl M. Oliner, 1988: *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: The Free Press.
- Roysten, Patrick, 2005a: Multiple imputation of missing values: update. *The Stata Journal* 5: 188-201.
- , 2005b: Multiple imputation of missing values: Update of ice. *The Stata Journal* 5: 527-536.
- Rubin, Donald B. und Neal Thomas, 1996: Matching Using Estimated Propensity Scores: Relating Theory to Practice. *Biometrics* 52: 249-264.
- StataCorp., 2007: *Stata user's guide. Release 10*. College Station, TX: Stata Press LP.
- Van Buuren, Stef, Hendrick C. Boshuizen und Dick L. Knook, 1999: Multiple Imputation of Missing Blood Pressure Covariates in Survival Analysis. *Statistics in Medicine* 18: 681-694.

Die Erklärung sozialen Handelns
Grundlagen und Anwendung einer integrativen Theorie
Kroneberg, C.
2011, X, 348 S., Softcover
ISBN: 978-3-531-17389-4