> „Der Vergleich ermöglicht den Blick über den nationalen Gartenzaun; er bietet die Chance, von anderen politischen Systemen zu lernen; er eignet sich für die Erfassung und Erklärung von Gemeinsamkeiten und Unterschieden; er schafft die Voraussetzungen für quasi-experimentelle Untersuchungsbedingungen und bildet die Grundlage erfahrungswissenschaftlich überprüfter Verallgemeinerung." (Schmidt 1995:348f)

Das für die empirische Überprüfung des zweistufigen Erklärungsansatzes und der Hypothesen angewandte Forschungsdesign beruht auf der Vergleichenden Methode und wird hier hinsichtlich von Fallauswahl, Vergleichsebene, Zeitraum und Messtaktung, Aggregationshöhe, Variablenauswahl, Datenbasis, Testverfahren und Ablauf dargestellt und diskutiert.

Fallauswahl: Entsprechend dem Most Similar Cases Design (Przeworski & Teune 1970) sollen die Geburtenentwicklung und ihre Bestimmungsfaktoren in den hoch entwickelten westlichen Demokratien, die OECD-Mitglieder sind, untersucht werden. Diese Länder erfüllen mit wenigen Ausnahmen gleichzeitig auch das zweite Kriterium, vom Phänomen des Zweiten Geburtenrückgangs betroffen zu sein. Hinsichtlich der Erfüllung dieser beiden Kriterien ist eine Vollerhebung anvisiert. Später der OECD beigetretene Länder wie Mexiko und Türkei werden nicht berücksichtigt, da sie während des Untersuchungszeitraums nicht vom Zweiten, sondern noch vom Ersten Geburtenrückgang betroffen waren. Kleinere Staaten wie Island und Luxemburg werden einbezogen, da von einer Wirkung der nationalen Rahmenbedingungen auf die Geburtenrate ausgegangen werden kann.[183] Insgesamt beläuft sich die Fallzahl auf 28 Länder (siehe 1.1), wobei es in zwei Fällen Umbrüche gibt, die im Zeitreihenvergleich berücksichtigt werden müssen: Beim Fall Deutschland werden vor 1990 Daten der BRD und danach die des vereinigten Deutschlands berücksichtigt[184], während bei Tschechien und der Slowakei auch vor 1993 getrennte Daten vorliegen. Die Auswahl der Fallzahl ist eine zentrale Weichenstellung des Forschungsdesigns

183 Bei den Jackknife-Analysen wird darauf geachtet, dass gemessene Zusammenhänge nicht nur auf Effekte dieser kleinen Staaten beruhen. Kritisch zum Problem stratifizierter Stichproben siehe: Ebbinghaus 2009:198f.

184 In den OECD-Datenbanken (u. a. OECD 2008a, 2009a) beziehen sich die Daten der Rubrik „Population" (u. a. TFR) auch vor 1990 auf Gesamtdeutschland, während sich andere Rubriken in diesen Zeitraum nur auf die BRD beziehen. Streng genommen handelt es sich um zwei unterschiedliche Fälle. Bei den multivariaten Querschnittsanalysen ist der Umbruch methodisch weniger bedeutend, wenn man sich den unterschiedlichen Zuschnitt der Fälle klar macht. Der Problematik dieser Umbrüche wird bei Veränderungsratenanalysen, die über das Jahr 1990 hinweg gehen, durch den Einbezug von Jackknife-Analysen Rechnung getragen. Darüber hinaus sind gravierende Auswirkungen nicht zu erwarten, da die Bevölkerungszahl der BRD erheblich größer als die der ehemaligen DDR ist und der gesamtdeutsche Wert zu rund 80% vom westdeutschen Wert determiniert ist.

https://doi.org/10.5771/9783845231211-214 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

und mit einem Zielkonflikt verbunden: Studien mit kleinem N oder gar Fallstudien ermöglichen einen tiefenschärferen Einblick in den Länderkontext, während eine hohe Fallzahl für die Anwendung statistischer Methoden und ihrem wissenschaftlichen Generalisierungspotenzial, aber auch für die Berücksichtigung vieler Variablen vorteilhaft ist (vgl. u. a. Berg-Schlosser 1997, Ebbinghaus 2009, Gschwend & Schimmelpfennig 2007, Lijphart 1971, 1975, Ragin 2000, Schmidt 1995).[185] Die hier vorgenommene Vollerhebung strebt eine größtmögliche Fallzahl innerhalb des gewählten Most Similar Cases Designs an, um die Voraussetzungen für statistische Verfahren zu erfüllen und zu verallgemeinbaren Aussagen zu gelangen. Die Freiheitsgrade sind bei N=28 mittelhoch und ermöglichen die Anwendung der multivariaten Regressionsmethode, insbesondere die gleichzeitige Inklusion mehrerer erklärender Variablen.[186] Gleichzeitig reduziert die Fallauswahl auch den Selection Bias (vgl. Jahn 2005), d. h. Verzerrungseffekte nicht repräsentativer Stichproben, da fast alle Länder, die vom Zweiten Geburtenrückgang vor 1990 betroffen sind, berücksichtigt werden. Diese Vorteile sind jedoch mit dem Nachteil erkauft, dass Aspekte des Länderkontextes nicht in der Tiefenschärfe von Studien mit kleinem N abgebildet werden können. Die skizzierten Vorteile überwiegen jedoch diesen Nachteil. Und als Studie mit mittlerem N ist das Forschungsdesign nicht nur variablenorientiert, sondern kann auch – insbesondere in der Residuendiagnostik – fallorientierte Aspekte aufnehmen (vgl. Ragin 2000:35). Ein Grundproblem vieler internationalen Vergleichsstudien, die Inklusion möglichst vieler Variablen bei begrenzter Fallzahl, ist auch in dieser Studie evident. Um dem zu begegnen, wird ein mehrstufiges Verfahren

185 Der Übergang zwischen hohem und niedrigem N ist graduell, Quantifizierungen als hoch, niedrig (oder mittel) selten und nicht einheitlich. Lijphart (1971:686) nennt als Beispiel für die „komparative Methode", die er gegenüber der „statistischen Methode" abgrenzt, weniger als zehn Fälle. (Hohe Fallzahl und statistische Methode sind nicht zwingend gleichzusetzen.) Ragin (2000:25) interpretiert Fallzahlen zwischen 3 und 50 als mittlere, Berg-Schlosser (1997:67) nennt dafür die Spanne 3 bis 20. Widmaier (1997:113) benennt als Grenze zwischen kleinen und großen Fallzahlen 30. Eine tiefenscharfe Kenntnis von mehr als 10 Untersuchungsfällen ist schwer zu leisten, anderseits sind für makroquantitative Querschnittsanalysen deutlich höhere Fallzahlen erforderlich (wobei PTS auch bei mittleren N durch Multiplikation von N und T (Zeit) eine ausreichende Fallzahl ermöglichen; siehe 6.1).

186 Schmidt (1982:130) weist in Anbetracht von 21 Fällen auf Begrenzungen für multivariate Auswertungsstrategien hin. Wagschal (1999) verweist auf zwei statistische Faustregeln, wonach mindestens 10 Freiheitsgrade insgesamt bzw. pro unabhängige Variable sinnvoll sind. Hier wird nur die erstere erfüllt. Derartige Faustregeln werden hier so aufgefasst, dass sie keine Fallbeilfunktion haben, sondern vielmehr hilfreiche Richtgrößen im Planungsprozess eines Forschungsvorhabens sind, um die Wahrscheinlichkeit robuster Ergebnisse zu erhöhen. Die Problematik niedriger Freiheitsgrade hinsichtlich Signifikanz (für N=20 vgl. Ragin 2000:27) und Robustheit ist dem Verfasser bewusst. Die multivariaten Ergebnissen der Kapitel 7 und 8 erweisen sich allerdings bei N=28 und vier unabhängigen Variablen als (hoch-)signifikant und robust.

https://doi.org/10.5771/9783845231211-214 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

mit bi- und multivariaten Analysen und Interkorrelationsanalysen angewendet (vgl. Schmidt 1982, 1995).

Vergleichsebene: Es wird unterstellt, dass das generative Verhalten in einem Land von den dort vorliegenden Rahmenbedingungen geprägt wird und dass durch ihre historische Entwicklung, gemeinsame Institutionen, Sprache und mediale Öffentlichkeit eine gewisse Homogenität dieser Bedingungen gegeben ist (vgl. Lesthaeghe 1995:34ff). Dabei ist zu beachten, dass in einigen Ländern sprachlich und kulturell eigenständige Minderheiten möglicherweise abweichende Geburtenraten haben (vgl. H-8d). Während die Kritik des methodologischen Nationalismus (Zürn 2001) auf die internationalen Zwänge hinweist, ist Kritik an der Analyseebene auch aufgrund des Whole-Nation Bias (Rokkan 1970:49, vgl. Lijphart 1975:166ff)[187] berechtigt, da regionale Differenzen innerhalb eines Landes nicht berücksichtigt sind. Dies betrifft in besonderem Maße die föderalen Staaten Deutschland, Schweiz oder USA (vgl. Schmidt et al. 2006). Der Einbezug föderaler Unterschiede würde den ohnehin enorm umfangreichen Datenumfang überfrachten (vgl. BMFSFJ 2007, Bonoli 2008, Höhn & Schubnell 1986:31). Da erhebliche Unterschiede durch nationalstaatlich variierende Faktoren erklärbar sind, ist die Auswahl der Analyseebene von Ländern nicht nur pragmatisch, sondern auch inhaltlich plausibel.

Eine methodologische Herausforderung für ländervergleichende Untersuchungsdesigns stellen Diffusionsprozesse dar. Das „Galton-Problem“ (vgl. Jahn 2009a) entsteht, wenn die wechselseitige Unabhängigkeit zwischen den untersuchten Fällen nicht gegeben ist und die beobachtbare Varianz dadurch beeinträchtigt wird. Umgekehrt – und gleichzeitig eine mögliche Lösung des Problems – können Diffusionsprozesse als erklärende Variable fungieren (vgl. Alber 1982). Die Analyse der Standardabweichung (siehe 1.1) zeigt eine Konvergenz der TFR bis Anfang der 1980er Jahre, was auf mögliche Diffusionsprozesse in dieser Phase hinweist. Methodisch wird von den Variablen Verhütungsmittelzugang und kultureller Wandel eine Schlüsselrolle für die internationale Diffusion angenommen. Durch die Berücksichtigung dieser Variablen im Zusammenspiel mit anderen in multivariaten Regressionsanalysen, der Identifizierung von Schub- und Bremskräften im Zeitverlauf und der Analyse zu mehreren Zeitpunkten (vgl. Jahn 2009a:98ff) wird der Diffusionsproblematik Rechnung getragen. Neben dieser internationalen Diffusion wird in dieser Untersuchung der Diffusionsgeschwindigkeit innerhalb der Länder ein zentraler Stellenwert zugeschrieben und ihre Determinanten überwiegend auf nationalstaatlicher Ebene vermutet.

Zeitraum und Messtaktung: Als Untersuchungszeitraum ist mit der Epoche von 1970 bis 2006 ein sehr langer Zeitraum gewählt. Damit soll ermöglicht wer-

187 Lijphart schätzt den Whole-Nation Bias als größte Schwäche der statistischen Methode auf Makroebene ein.

https://doi.org/10.5771/9783845231211-214 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

den, sowohl den Beginn des Geburtenrückgangs als auch bis in die Gegenwart reichende Entwicklungen erklären zu können. Da Querschnittsmodelle mit Niveaudaten „dynamisch unterspezifiziert" (Widmaier 1997:116) sind, wird neben der synchronen auch die diachrone Betrachtung vorgenommen. Letztere wird durch die Analyse von Veränderungsraten durchgeführt. Diese werden für mehrere Zeitabschnitte vorgenommen, wobei die Auswahl der Subperioden systematisch mit dem zweistufigen Erklärungsansatz und den bivariaten Voruntersuchungen verknüpft ist. Der Vergleich wird hinsichtlich der Zeitdimension dadurch in doppelter Weise sehr breit angelegt. Diese Breite ist bedeutsam, da der Verfasser davon ausgeht, dass die gezeigten Zusammenhänge zeitlich nicht invariat sind, d. h. ein im Querschnitt eines Jahres gezeigter Befund nicht als Erklärung für die gesamte Epoche des Geburtenrückgangs aufgefasst werden kann. Vielmehr ergibt die systematische Verknüpfung mehrerer engmaschiger Niveau- und Veränderungsratenanalysen über den Zeitverlauf hinweg die Möglichkeit, ein dynamisches Erklärungsmuster des Phänomens zu identifizieren.

Aggregationshöhe: Die empirische Basis dieser Untersuchung bilden Aggregatdaten. Ein Einbeziehen der Mikroebene in das Untersuchungsdesign hätte eine schlechte Kosten-Nutzen-Relation. Zum einen ist eine umfassende Beschaffung der relevanten Individualdaten für den großen Untersuchungszeitraum und die Fallzahl kaum möglich. Zu beachten ist dabei, dass ein schlicht auf ökonomischen Faktoren basierendes Mikromodell zu viele entscheidende Variablen übersehen würde. Zum anderen würde die alleinige Analyse der Mikrodimension nicht die Erkenntnisgewinne bringen, die für eine Beantwortung der Forschungsfragen nötig sind.[188] Die Studie soll politische, ökonomische, technologische und gesellschaftliche Rahmenbedingungen im internationalen Vergleich analysieren und mit deren zeitlichen und räumlichen Variationen ein Licht auf die Ursachen der Entwicklungen der Geburtenraten werfen. Dies kann mit hoch aggregierten Makrodaten am besten gelingen. Grundlage dieser Erwartung ist die plausible Annahme, dass das durchschnittliche generative Verhalten einer Gesellschaft maßgeblich von den genannten Rahmenbedingungen geformt wird.[189]

Allerdings ist es hilfreich – wenn nicht gar unabdingbar – darüber hinaus die theoretische Verbindung zur Mikroebene herzustellen (vgl. Kittel 2006:654).

188 Vgl. Kaufmann (2005:18) zu einer ähnlichen Fragestellung: "(Es) würde die zusätzliche Einbeziehung der Mikrodimension m. E. der Klarheit nicht förderlich sein."

189 Lesenswerte Begründungen für die Analyse von Makrodaten bezogen auf die Geburtenrate siehe Castles (1998:264), Lesthaeghe (1995:34) und in anderem Zusammenhang Schmidt (1993a:19, 1995:334). Es gibt jedoch auch skeptische Positionen gegenüber der Aussagekraft von Aggregatdatenanalysen (u. a. Hoem 2008:253; Neyer & Andersson 2008:707). Gauthier (2007:331) betont für Aggregatdaten die Notwendigkeit multivariater Untersuchungsdesigns, da die häufig zitierten bivariaten Querschnittskorrelationen den Einfluss anderer Determinanten nicht erfassen.

https://doi.org/10.5771/9783845231211-214 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Deswegen wurde in Kapitel 3 die potenzielle Wirkungsweise der einzelnen Makrofaktoren in dem generativen Mehrebenenmodell mit dem Mikromodell der Fertilitätsentscheidung in Bezug gesetzt und plausibilisiert. Bei der Diskussion der empirischen Ergebnisse werden diese ebenfalls anhand des Mikromodells der Fertilität auf der Individualebene nachvollzogen (siehe 6.5 und 7.4). Das Aufgreifen der Mikroebene soll der Vermeidung eines ökologischen Fehlschlusses[190] dienen und die Rückkopplung der Ergebnisse nicht nur zu den bestehenden theoretischen Makroansätzen, sondern auch zu Ansätzen auf der Mikroebene gewährleisten. Neben dieser theoretischen Rückkopplung ist eine zusätzliche empirische Überprüfung der Ergebnisse anhand von Individualdatenanalysen zweifellos sinnvoll, diese kann in dieser Untersuchung jedoch nicht geleistet werden.

Variablenauswahl: Die zu erklärende Größe dieser Untersuchung, die abhängige Variable, ist die TFR (zur Diskussion alternativer Indikatoren vgl. 1.1) und ihre Veränderungsrate. Als unabhängige Variablen werden Indikatoren verwendet, die die im vorigen Abschnitt entwickelten Hypothesen überprüfen. Hinsichtlich der Variablenauswahl ist ein Y-zentriertes Forschungsdesign[191] (vgl. Ganghof 2005, Gschwend & Schimmelpfennig 2007) zu konstatieren, das mit der Forschungsfrage begründet wird. Es wird folglich versucht, die Varianz der abhängigen Variablen, der Geburtenrate, durch die Einbeziehung möglichst aller potenziell relevanten unabhängigen Variablen zu erklären. Der Umgang mit einer großen Zahl unabhängiger Variablen im multivariaten Teil erfordert neben bivariaten Voruntersuchungen auch theoretische Überlegungen, um die Zahl der Regressionsrechnungen halbwegs überschaubar zu halten (vgl. Sala-I-Martin 1997) und den Freiheitsgraden der Fallauswahl Rechnung zu tragen (vgl. 6.1).

Datenbasis: Für die Datenbasis wird v. a. auf verschiedene OECD-Datenbanken zurückgegriffen (OECD 1999, 2001e, 2007b, 2008a, 2008b, 2009a, 2009b, 2009c), die teilweise mit Daten aus offiziellen nationalen Ämtern (u. a. Statistisches Bundesamt 2008c) ergänzt sind. Daneben werden Daten von Barrett et al. (1982, 2001), Castles (1998), Freedom House (2009), Inter-Parliamentary Union (2009), Luxembourg Income Study (2009), Schmidt (2009), United Nations (2002, 2003b, 2006, 2007a) und World Values Surveys (2006, vgl. Pickel 2009) verwendet. Für einige Länder, insbesondere die neuen osteuropäischen Demokratien, sind nur vergleichbare Daten aus der zweiten Hälfte des Untersu-

190 Dazu sind auch Aspekte des Forschungsdesigns wie die Analyse von Niveau- und Veränderungsdaten hilfreich.

191 Auch wenn der Schwerpunkt auf Y liegt, werden in der Analyse auch X-Elemente betrachtet und neue Erkenntnisse hinsichtlich ihrer Verknüpfung bzw. ihres Zusammenspiels versucht zu generieren. Der Y-Schwerpunkt eines Forschungsdesigns muss nicht zwingend mit einer X-Blindheit einhergehen – und umgekehrt.

https://doi.org/10.5771/9783845231211-214 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

chungszeitraums verfügbar; dem wird durch eine Reduzierung des Länderpools um diese Staaten vor 1990 Rechnung getragen.

Testverfahren und Ablauf: Das Forschungsdesign ist mehrstufig. In Kapitel 5 wird der Zusammenhang der TFR mit den einzelnen unabhängigen Variablen anhand bivariater Korrelationen vor dem Hintergrund der Interkorrelationsmatrixen systematisch betrachtet. Die Korrelationsentwicklung wird auch im Zeitverlauf analysiert. Hierbei ist von besonderem Interesse, bei welchen Variablen und in welchem Jahr sich der Richtungszusammenhang innerhalb des Untersuchungszeitraums dreht. Diesen Erkenntnissen folgt eine Aufteilung des Geburtenrückgangs in Phasen, die für die Analyse der Veränderungsraten maßgeblich ist. Die multivariaten Regressionsanalysen[192], die engmaschig Niveau- und Veränderungsdaten für mehrere Jahre bzw. Zeitspannen umfassen, werden in Kapitel 6 für die Diffusionsphase und in Kapitel 7 für die Akkomodationsphase dargestellt. Dabei werden die Effekte der einzelnen Determinanten durch systematische Einbeziehung der anderen Determinanten kontrolliert und Erklärungsmodelle mit hoher potenziell erklärter Varianz generiert. Die im Zeitverlauf variierenden Regressionsbefunde werden in Bezug zueinander und zum Erklärungsansatz interpretiert und fließen in ein differenziertes Gesamterklärungsmodell ein.

192 Die Voraussetzungen und Probleme der Regressionsmethode sowie der in dieser Untersuchung praktizierte Umgang mit ihnen werden in Abschnitt 6.1 ausführlich diskutiert.

https://doi.org/10.5771/9783845231211-214 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY