

KAPITEL 2

Methodische Grundlagen der Untersuchung

Simulationen sind in der Politikwissenschaft noch ein junges Feld (vgl. Axelrod 2007: 90; Saam 2015: 2). Insbesondere in der deutschen Politikwissenschaft wurden sie als Methode lange Zeit nur selten angewandt. Gleichwohl haben Simulationen im Zuge der „Verwissenschaftlichung“ (Kittel 2009: 580) der Politikwissenschaft und dem damit einhergehenden größeren Stellenwert von methodologischen Fragen im politikwissenschaftlichen Diskurs international – besonders im angelsächsischen Raum – in den letzten Jahren zunehmend an Bedeutung zur Generierung und Überprüfung von Theorien gewonnen (vgl. Garson 2009: 267; Kittel 2009: 578; Saam 2015: 2–3). Die Mehrheit dieser formalen Modelle basiert auf *Rational Choice*-Annahmen, die für Politikwissenschaftler von großem Interesse sind, da sie eine mathematische Analyse strategischer Interaktionen zwischen Akteuren bzw. Akteursgruppen ermöglichen (vgl. Axelrod 2007: 93; Morton 1999: 82). Daneben wurde in jüngster Zeit eine steigende Zahl an agentenbasierten Modellen entwickelt (vgl. Flache/Mäs 2015; Marchi/Page 2014), die auch in der deutschen Politikwissenschaft auf wachsendes Interesse stoßen (siehe z. B. Behnke 2009: 195; Brunner 2012; vgl. Lorenz 2012; Shikano 2008).

Im Folgenden werden zunächst die methodischen und wissenschaftstheoretischen Grundlagen von Simulationen skizziert (Abschnitt 2.1). Es soll erläutert werden, auf welchen Annahmen Simulationen beruhen, wie sie im Forschungsprozess eingesetzt werden können, welche Form von Erkenntnis sie generieren können und wo ihre Grenzen als methodische Instrumente liegen. Nach dieser Einführung werden Forschungsdesign und -ergebnisse des Projekts „The European Union Decides“ (Thomson et al. 2006b) sowie die daran anschließende Studie „Resolving Controversy

in the European Union“ (Thomson 2011) diskutiert, die Entscheidungsprozesse in der EU mittels formaler Modelle untersuchen. Sie liefern Erkenntnisse darüber, wie hoch die Prognosefähigkeit verschiedener Modelle bei der Nachbildung von Ratsentscheidungen in der EU ist und unter welchen Kontextbedingungen die Modelle am besten funktionieren (vgl. Thomson/Hosli 2006: 6–7). Die vorliegende Arbeit baut mit der Anwendung eines auf den DEU-Datensatz getesteten Modells auf diese Forschungsergebnisse auf und gliedert sich in den Forschungszweig ein. Zur weiteren Konkretisierung wird in Abschnitt 2.3 das in dieser Arbeit angewandte Modell, das *predictioneer's game* von Bruce Bueno de Mesquita, sowie dessen Anpassung an die europäische Erdgasaußenpolitik erläutert.

2.1 Simulationen in der Politikwissenschaft

Differenzierung zwischen beobachtbarer Wirklichkeit und der ihr zugrundeliegenden Kausalstruktur

Simulationen basieren auf der Annahme, dass politische Ereignisse, die wir in der Wirklichkeit beobachten, nicht einem Zufall entspringen, sondern aus bestimmten Ursachen resultieren (vgl. Behnke 2009: 176). Behnke spricht in diesem Zusammenhang von einem „Daten generierenden Prozess“ [im Folgenden: DGP] (Behnke 2009: 176; siehe auch Morton 1999: 33–34). Sie beziehen sich damit auf die der Wirklichkeit zugrundeliegende und die Wirklichkeit entsprechend evozierende Struktur: „Dieser Daten generierende Prozess ist verantwortlich für alle beobachtbaren Phänomene, er entspricht der Kausalstruktur, nach der sich die Wirklichkeit gestaltet und gestalten muss.“ (Behnke 2009: 176; die Hervorhebung entspricht dem Originaltext) Zum besseren Verständnis einer solchen Kausalstruktur hilft die Analogie zur Physik: Blätter an Laubbäumen erscheinen uns gewöhnlich grün. Der Grund hierfür ist, dass das Chlorophyll (v.a. die Typen a und b) in den Blättern aus dem kontinuierlichen Farbspektrum des Sonnenlichts bevorzugt die Farben Blau und Rot absorbiert und die Blätter somit Licht mit der Restfarbe Grün in unser Auge streuen. Quantenphysikalisch betrachtet lässt sich dieses Phänomen so erklären, dass Chlorophyll-Moleküle im Grundzustand unbesetzte Energieniveaus aufweisen, die eine Absorption von Lichtquanten im Wellenlängenbereich 400 – 500 nm (violett – blau) und im Bereich 600 – 700 nm (orange – rot) zulassen. Die angeregten Zustände der Moleküle sind nur von kurzer Dauer (10^{-9} bis 10^{-12} Sekunden), die Elektronen kehren somit wieder in den Grundzustand zurück und geben dabei die aufgenommene Energie

wieder ab; diesmal aber nicht als Licht, sondern vor allem durch Energietransport an benachbarte Moleküle, was letztendlich ein Mechanismus zur Energieaufnahme der Pflanze ist. Aus diesem Grund fehlen die genannten Wellenlängen im Spektrum des Lichts, das von den Blättern gestreut wird: Man sieht grüne Blätter (vgl. Bader 2000: 270).

Dieses Beispiel verdeutlicht, wie kompliziert der Rückschluss von einem beobachtbaren Phänomen auf die dieses Phänomen bewirkenden Gesetze, den DGP, ausgestaltet sein kann, da das physikalische Gesetz selbst nicht direkt beobachtbar ist. Unmittelbar sichtbar sind lediglich die von ihm erzeugten Effekte, was die Erforschung des DGP maßgeblich erschwert (vgl. Behnke 2009: 176; Kroneberg/Kalter 2012: 74). Auch für die politische Sphäre wird ein solcher Zusammenhang von DGP und beobachtbarer Wirklichkeit angenommen: Viele politischen Phänomene, die wir beobachten können, basieren auf bestimmten Gesetzmäßigkeiten, die selbst wiederum nicht direkt sichtbar sind. Um die jeweiligen politischen Phänomene fundiert zu analysieren, muss diese Kausalstruktur daher aufgedeckt werden. Mittels Simulationen geschieht dies auf indirektem Wege. Sie dienen dazu, die zu erklärenden Regelmäßigkeiten nachzubilden und durch Rückschlüsse von der Simulation auf die tatsächliche Wirklichkeit zu erforschen: „Wir können uns [...] der Erforschung des DGP nur auf indirekte Weise annähern. Die einzige Vorgehensweise, die uns überhaupt möglich ist, besteht darin, ein *Modell* des DGP zu ‚basteln‘, das den DGP in seinen wichtigsten strukturellen Eigenschaften möglichst genau abbilden soll.“ (Behnke 2009: 176; siehe auch Dehling/Schubert 2011: 43–46; Morton 1999: 62) Daraus ergibt sich der Forschungszweck, der durch den Einsatz von Simulationen verfolgt wird: Ziel ist es nicht, singuläre Ereignisse zu untersuchen, sondern die Konstanz bestimmter politischer Phänomene, d.h. die ihnen zugrunde liegende Struktur zu analysieren (vgl. Behnke 2009: 175–176).

Im Folgenden wird genauer dargelegt, wie Simulationen als virtuelle Welten zur Analyse politischer Phänomene und der ihnen zugrundeliegenden Wirkungsmechanismen im Forschungsprozess eingesetzt werden können.

Virtuelle Welten als Analyseinstrumente für die Realität

Bei der Entwicklung von Simulationen wird auf der Grundlage von explizit geäußerten Annahmen über den zu erforschenden Gegenstand eine virtuelle Wirklichkeit erstellt, die eine möglichst große Ähnlichkeit zur tatsächlichen Wirklichkeit besitzen soll: „The core idea of simulation is to create a “virtual world” of the simulation

that is somehow an analogy to the “real world”.“ (Hegelich 2016: 455; siehe auch Behnke 2009: 175) Schließlich wird mit dem Einsatz von Simulationen in der Wissenschaft das Ziel verfolgt, durch die Analyse der virtuellen Welt Rückschlüsse auf die wirkliche Welt ziehen zu können. Dazu müssen jedoch die Regeln von virtueller Welt und wirklicher Welt, d.h. des DGPs, weitgehend übereinstimmen – nur dann können begründete Hypothesen über die Gesetzmäßigkeiten und die beobachtbaren Phänomene der wirklichen Welt aufgestellt werden: „You can formulate expectations about other systems – including real world systems – on the basis of *ceteris paribus* assumptions: To the extent the same rules apply to the simulation and to the compared system, both systems should behave similar.“ (Hegelich 2016: 456)²⁷

Da die der Wirklichkeit zugrundeliegende Kausalstruktur nicht direkt beobachtbar ist, ist ein direkter Vergleich der Regeln von virtueller und wirklicher Welt allerdings nicht möglich. Was jedoch beobachtbar ist, sind die politischen Phänomene, die der DGP in der Wirklichkeit erzeugt. Um herauszufinden, ob eine Simulation die Struktur der tatsächlichen Wirklichkeit möglichst genau abbildet, also eine Analogie zwischen virtueller und wirklicher Welt besteht, werden daher die Ergebnisse, die die beiden Welten hervorbringen,²⁸ verglichen. Die von Hegelich (2016: 456) angeführte Erläuterung, dass virtuelle und wirkliche Welt sich gleich verhielten, sofern sie auf den gleichen Regeln beruhten, wird somit umgedreht: Generieren virtuelle Welt und wirkliche Welt (nahezu) identische Ergebnisse, ist zu vermuten, dass die Regeln, die diese Phänomene hervorbringen, ebenfalls übereinstimmen. Entsprechend erläutert Behnke:

Die Konfrontation unseres Modells der Realität mit derselben beziehungsweise dem DGP findet also niemals unmittelbar statt, sondern indirekt, indem wir die logischen Konsequenzen aus dem Modell in Form von Beobachtungssätzen

27 *Ceteris paribus* bedeutet „unter sonst gleichen Bedingungen“, d.h. es wird ausschließlich der Effekt eines zu überprüfenden Faktors auf das Simulationsergebnis gemessen. Alle anderen Aspekte bleiben konstant (vgl. Morton 1999: 38–39).

28 Gemeint sind hier als Ergebnisse der virtuellen Welt die Simulationsergebnisse, die aus den zuvor aufgestellten Annahmen resultieren. In der Spieltheorie werden diese als Gleichgewichte bezeichnet. Generell herrscht jedoch keine Einheitlichkeit für die Lösungskonzepte von formalen Modellen vor: „After the situation is transformed into a formal set of assumptions expressed in abstract terms, the system is then studied or “solved” for predictions, which are presented as theorems, propositions, or just “results”. Solving a model means exploring the implications of the set of assumptions. Solution concepts typically depend on the formulation of the model.“ (Morton 1999, 49)

mit den sich zwingend ergebenden Konsequenzen aus dem DGP in Form von Beobachtungen vergleichen. Entsprechen die vorhergesagten Beobachtungssätze beziehungsweise Basissätze weitgehend den tatsächlich gemachten Beobachtungen, dann schließen wir daraus indirekt (nicht logisch), dass unser Modell des DGP zumindest ein brauchbares Abbild des tatsächlichen DGP zu sein scheint. (Behnke 2009: 177)

Simulationen werden somit auf der Grundlage von Erfahrungen mit der wirklichen Welt erstellt und die Analogie der beiden Systeme durch einen Vergleich des Simulationsergebnisses mit den beobachtbaren Phänomenen der tatsächlichen Wirklichkeit überprüft. Auf diese Weise können Simulationen dazu eingesetzt werden, bestimmte Segmente von Realität bzw. die ihr zugrundeliegenden Gesetzmäßigkeiten und kausalen Zusammenhänge nachzuvollziehen und zu analysieren (vgl. Garson 2009: 274). Zudem ist es möglich, die Passgenauigkeit von Ergebnissen verschiedener Simulationen zu vergleichen und daraus zu schließen, welche Simulation bzw. welche Kombination von Simulationen am besten dazu geeignet ist, die Kausalstruktur der Wirklichkeit abzubilden (vgl. Morton 1999: 55). Simulationen dienen somit als methodisches Instrument zur Generierung und Überprüfung von Theorien über die Wirklichkeit (vgl. Bechtel/Leuffen 2010: 310; Behnke 2009: 192; Hegelich 2016: 462; Ray/Russett 1996: 446).

Aus den bisherigen Erläuterungen lässt sich schließen, dass der wissenschaftliche Einsatz von Simulationen zur Analyse von Wirklichkeit ein Spannungsfeld zwischen virtueller und wirklicher Welt impliziert, das Konsequenzen für die Beweiskraft von analytischen Erkenntnissen durch den Simulationsansatz hat: Wenngleich versucht wird, ein möglichst adäquates Abbild von der Wirklichkeit zu kreieren, um Rückschlüsse auf ebendiese ziehen zu können, bleiben Simulationen doch stets artifizielle Konstrukte. Schließlich sind Simulationen notwendigerweise immer durch eine Reduktion von Komplexität gekennzeichnet. Eine vollständige Reproduktion der Wirklichkeit ist nicht möglich – sonst würde es sich um ein Experiment handeln (vgl. Behnke 2009: 175; Hegelich 2016: 456). Daher ist zu betonen, dass die Unterscheidung von virtueller und wirklicher Welt in der Analyse stets gewahrt werden muss, nicht zuletzt bei der Interpretation der Forschungsergebnisse (vgl. Johnson 1999: 1518, 1524). Schlussfolgerungen von Erkenntnissen über die Simulation auf die Wirklichkeit können aufgrund dieser Trennung von virtueller und wirklicher Welt nicht verifiziert werden. Simulationen liefern daher keine Beweise für kausale Erklärungen über die Wirklichkeit (vgl. Hegelich 2016: 470; siehe auch Dehling/Schubert

2011: 45). Damit sollen die zuvor aufgeführten Erläuterungen zum Erkenntnisgewinn durch Simulationen in der Politikwissenschaft nicht negiert, sondern lediglich die Grenzen von Simulationen im Forschungsprozess gekennzeichnet und auf diese Weise dem häufig begangenen Fehler einer Verwechslung von Simulation und Wirklichkeit vorgebeugt werden:

[... T]here is always a problem with simulations when it comes to causalities. However – as will be seen – this problem is far from unique to the simulation approach. Sometimes, scientists using a simulation approach are so convinced of their model that they forget the difference between virtual and real world. They might claim that their simulation has ‘proved’ some real world effect. But because of *ceteris paribus* the connection to the real world is always questionable. Overestimating the relevance of one owns simulations will make you ‘easy prey’ for any critiques. (Hegelich 2016: 456)

Simulationen als Ansatz zwischen Induktion und Deduktion

Simulationen können keinem der beiden standardmäßigen Verfahren zur Gewinnung wissenschaftlicher Erkenntnisse – weder der Induktion, noch der Deduktion – eindeutig zugeordnet werden. Behnke (2009, 175–176) beschreibt die Entwicklung und Anwendung von Simulationen als eine Verbindung von Induktion und Deduktion: Die Konstruktion der Simulation in Anlehnung an die bereits gemachten Beobachtungen der Wirklichkeit kennzeichnet er als induktiven Bestandteil des Forschungsprozesses. Die Ergebnisse bzw. Schlussfolgerungen, die die Simulation aus den zuvor aufgestellten Annahmen generiert und dem Wirklichkeitssegment, das simuliert wird, zugeschrieben werden, charakterisiert er als deduktive Komponente des Forschungsprozesses. Es ist jedoch fraglich, ob eine derartige Unterteilung des Forschungsprozesses sowie die Zuschreibung von induktiven und deduktiven Bestandteilen bei der Entwicklung und Auswertung von Simulationen sinnvoll sind, da sie den Anschein erwecken, die bewusste Trennung von virtueller und wirklicher Welt relativieren zu wollen. Axelrod grenzt den Simulationsansatz daher von Induktion und Deduktion ab und beschreibt ihn als einen neuen Weg der Ausübung von Wissenschaft:

Simulation as a way of doing science can be contrasted with the two standard methods of induction and deduction. Induction is the discovery of patterns in empirical data. [...] Deduction, on the other hand, involves specifying a set of axioms and proving consequences that can be derived from those assumptions. [...] Like deduction, it starts with a set of explicit assumptions.

But unlike deduction, it does not prove theorems. Instead, a simulation generates data that can be analyzed inductively. Unlike typical induction, however, the simulated data comes from a rigorously specified set of rules rather than direct measurement of the real world. While induction can be used to find patterns in data, and deduction can be used to find consequences of assumptions, simulation modeling can be used as an aid intuition. (Axelrod 2007: 92–93)

Der Nutzen von Simulationen als Hilfestellung für die Intuition bezieht sich nicht nur auf die Analyse gegenwärtiger bzw. vergangener politischer Sachverhalte. Gelingt es, mittels Simulationen die Kausalstruktur der Wirklichkeit aufzudecken, bzw. sich dieser möglichst anzunähern, ist es auf dieser Grundlage ebenfalls möglich, wissenschaftlich fundierte Hypothesen über *zukünftige* Ereignisse aufzustellen (vgl. Dehling/Schubert 2011: 45–46). Im Folgenden werden Möglichkeiten und Grenzen von Simulationen zur Generierung von Zukunftsprognosen dargelegt.

Simulationen als Instrument für Zukunftsprognosen

„Anticipating the future is both a social obligation and intellectual challenge that no scientific discipline can escape“, schrieben Gerald Schneider, Nils Petter Gleditsch und Sabine C. Carey (2010b: 1) im Jahr 2010, basierend auf einem Vortrag zur 50. jährlichen Tagung der International Studies Association in New York (15.-18. Februar 2009). Analog zu diesem Auftrag für die Internationalen Beziehungen betonten Bechtel und Leuffen, wie wichtig von der Politikwissenschaft erarbeitete Zukunftsprognosen für politische Entscheidungsträger sind: „Forecasts are indispensable for actors in the real world. In order to make informed decisions, political actors (legislators, bureaucrats, as well as citizens) need to have an idea of the consequences of their actions. Therefore, forecasting offers crucial information to anticipate, and if necessary counteract, important developments.“ (Bechtel/Leuffen 2010: 309–310; siehe auch Miller 2014b) Dennoch herrscht im politikwissenschaftlichen Methodendiskurs noch immer eine große Skepsis vor, ob Vorhersagen in der politischen Sphäre überhaupt möglich sind (siehe z. B. Stevens 2012; Zambarnardi 2016; für eine Skizzierung der Debatte und eine Diskussion methodischer Fortschritte in der Entwicklung von Vorhersagen siehe u.a. Armstrong et al. 2015; Gleditsch/Ward 2013; Schneider et al. 2011). In Publikationen über Simulationen in der Politikwissenschaft steht deren Nutzen für die Entwicklung von Zukunftsprognosen häufig im

Vordergrund (vgl. Axelrod 2007: 92).²⁹ Entsprechend sind Simulationsansätze eingebettet in die Auseinandersetzung um Möglichkeiten und Grenzen von Vorhersagen in der Politikwissenschaft. Betrachtet man die Vorgehensweise, wie Simulationen als Prognoseinstrumente eingesetzt werden können, so wird deutlich, dass sowohl Befürwortern als auch Kritikern von Zukunftsprognosen in gewisser Hinsicht Recht gegeben werden kann. Dies soll im Folgenden dargelegt werden.

Anknüpfend an die zuvor erläuterten Annahmen des Simulationsansatzes dienen Modelle dazu, die der Wirklichkeit zugrundeliegende Kausalstruktur zu emulieren. Sobald man ein Modell entwickelt hat, das die Wirklichkeit möglichst akkurat abbildet, die Regeln von virtueller und wirklicher Welt also miteinander vereinbar sind, ist davon auszugehen, dass auch die beobachtbaren Phänomene, die diese Regeln generieren, identisch sind (vgl. Hegelich 2016: 456). Folglich kann das Modell in einem nächsten Schritt dazu genutzt werden, wissenschaftlich fundierte Hypothesen über zukünftige Entwicklungen des jeweiligen simulierten Wirklichkeitssegments zu generieren (vgl. Axelrod 2007: 96; Johnson 1999: 1511). Auf der Grundlage von *ceteris paribus*-Annahmen wird getestet, welche Effekte die Variation einzelner Variablen auf das Simulationsergebnis haben und daraus abgeleitet, wie sich die tatsächliche Wirklichkeit – gegeben der Analogie von virtueller und wirklicher Welt und unter ansonsten gleichbleibenden Bedingungen – entwickeln würde, wenn sich in der Realität die gleichen Variablen verändern würden. Entwickelt man z. B. ein Modell, das die Kausalstruktur polnischer Energiepolitik speziell hinsichtlich der Frage des nationalen Energiemixes nachbildet, so können auf Grundlage der aus dem Modell gewonnenen Erkenntnisse Hypothesen darüber aufgestellt werden, wie sich die Förderung von Schiefergas in Polen auf die Energiepolitik und den Energiemix auswirken werden – unter der Voraussetzung, dass alle anderen Parameter unverändert

29 Damit soll nicht behauptet werden, Zukunftsprognosen wären der einzige Zweck des Simulationsansatzes. Vielmehr verweisen Wissenschaftler, die mit der Anwendung von Simulationen vertraut sind, stets auf den vielfältigen Nutzen, den Simulationen für die Politikwissenschaft entfalten können. So führt Epstein – um lediglich ein Beispiel zu nennen – sechzehn Gründe jenseits von Vorhersagen für die Anwendung formaler Modelle in der Forschung auf (2008; siehe u.a. auch Axelrod 2007, 2004; Hegelich 2016; Saam/Gautschi 2015). In der vorliegenden Arbeit liegt der Schwerpunkt jedoch auf der Generierung von Theorien über die Entscheidungsfindung in der europäischen Erdgasaußenpolitik sowie der Erstellung von Zukunftsszenarien für dieses Politikfeld. Entsprechend wird in diesem Kapitel ein Fokus auf die Entwicklung von Simulationen als Instrumente für Zukunftsprognosen gelegt.

bleiben. Auf diese Weise können in der Politikwissenschaft politische Entwicklungen prognostiziert sowie die Auswirkungen verschiedener Handlungsoptionen getestet und evaluiert werden.

Die methodische Vorgehensweise bei der Anwendung des Simulationsansatzes ist somit zunächst unabhängig davon, ob die Modellanalyse dazu genutzt werden soll, Rückschlüsse auf vergangene, gegenwärtige oder zukünftige politische Entwicklungen zu ziehen. Sie beginnt stets mit der Nachbildung des zu erklärenden Gegenstands und der damit verbundenen Aufdeckung der grundlegenden Wirkungsmechanismen. Darauf verweisen auch Bueno de Mesquita, Newman und Rabushka, um die Skepsis der Politikwissenschaft gegenüber der Generierung von Zukunftsszenarien zu relativieren und zugleich deren unbedingte Notwendigkeit zu betonen:

The twin objectives of science are explanation and prediction. Political science seeks to explain why people behave politically as they do, why political processes and institutions function as they do, and why specific outcomes occur. The logical structures of explanation and prediction are the same; the difference between them is only whether the scientist's objective is to account for a past event or describe a future event. Thus prediction is a major concern of political scientists, especially in cases where it is essential to anticipate the consequences of political actions in order to provide sound advice on policy matter. (Bueno de Mesquita et al. 1985: 6)

Gleichwohl ist die Kritik an der Einsatzfähigkeit von Simulationen für Prognosen insofern berechtigt, als zwei signifikante Beschränkungen der Prognosefähigkeit in einem solchen Forschungsansatz stets impliziert sind: Zum einen müssen auf der Grundlage von Simulationen entwickelte Zukunftsprognosen von einer *ceteris paribus*-Bedingung ausgehen, die Schlussfolgerungen von der virtuellen auf die wirkliche Welt stets relativiert. Daher sind Forschungsansätze, in denen eine unmittelbare Übertragung der aus der Simulation gewonnenen Erkenntnisse auf die Realität behauptet wird, stets problematisch. Zum anderen bringt der zuvor benannte Forschungszweck von Simulationen – die Enthüllung der Kausalstruktur, die der Wirklichkeit zugrundeliegt – mit sich, dass Simulationen keine singulären, sich nicht aus der Logik der Regeln ergebenden Ereignisse vorhersagen können (vgl. Hegelich 2016: 455–456). Mit der Mehrzahl der in der Politikwissenschaft entwickelten Modelle – und dies gilt auch für den Forschungszweig, an den die vorliegende Arbeit anknüpft – wird darauf abgezielt, die elementaren Mechanismen, d.h. bestimmte Muster politischer Prozesse zu identifizieren und nachzubilden, die

nicht nur für einen spezifischen, sondern für eine größere Anzahl von Fällen gelten (vgl. Thomson/Hosli 2006: 11). Daraus folgt, dass auch die Prognosen, die derlei Simulationen generieren, auf Muster und Tendenzen zukünftiger Entwicklungen beschränkt sind: „It is important to understand models of social and political phenomena are not likely to provide predictions of specific events such as the assassination that set off World War I, or the date at which the next Asian currency crisis will start. What good models can, however, provide are predictions of some pattern of events such as the dangerous instability among the Great power in 1914 [...]“. (Axelrod 2004: 2)³⁰ Dies sei am Beispiel der amerikanischen Sicherheitspolitik verdeutlicht: Als am 11. September 2001 ein Terroranschlag auf die Zwillingstürme des World Trade Centers verübt wurde, wäre es für politikwissenschaftliche Simulationen nahezu unmöglich gewesen, ein spezifisches Datum sowie die detaillierte Ausgestaltung der in der Folge von der USA geführten Intervention in Afghanistan zu prognostizieren. Was politikwissenschaftliche Simulationen aber leisten können, ist, die Gesetzmäßigkeiten vorheriger amerikanischer Sicherheitspolitik zu identifizieren und Hypothesen darüber zu entwickeln, wie sich dieses Politikfeld nach den Terroranschlägen tendenziell entwickeln wird.

Verbunden mit der Intention, Simulationen für Zukunftsprognosen einzusetzen, ist eine Debatte um die Zielkonflikte Simplizität und Exaktheit von Modellen. In Anlehnung an den Militärslogan „keep it simple, stupid“ dominiert insbesondere unter Modellierern von agentenbasierten Ansätzen das KISS-Prinzip, das vorgibt, die den Modellen zugrundeliegenden Annahmen möglichst einfach und nachvollziehbar zu gestalten (vgl. Axelrod 2007: 93). Begründet wird dies mit den begrenzten kognitiven Fähigkeiten der Menschen: „The KISS principle is vital because of the character of the research community. Both the researcher and the audience have limited cognitive ability. When a surprising result occurs, it is very helpful in giving other researchers a realistic chance of replicating one’s model and extending the work in new directions.“ (Axelrod 2007: 93) Sofern es also darum geht, grundsätzliche Zusammenhänge oder Prozesse zu verstehen,

30 An dieser Stelle sei zu unterstreichen, dass die hier aufgeführten Beschränkungen sich in erster Linie auf den in dieser Arbeit angewandten Modellansatz beziehen. Selbstverständlich werden auch Simulationen entwickelt, deren Zweck es ist, Extremereignisse o.ä. zu prognostizieren. Gerade in der Politikwissenschaft liegt der Schwerpunkt von Simulationen aber auf der Analyse und Prognose genereller Mechanismen und Tendenzen politischer Prozesse.

ist die Simplität und Nachvollziehbarkeit der Modellregeln gemeinhin wichtiger als die hochgradig detaillierte Darstellung einer bestimmten Situation oder eines bestimmten Gegenstands. Die Erfüllung dieser Prämisse wird jedoch erschwert, sobald Simulationen für Zukunftsprognosen eingesetzt werden. Schließlich ist es in diesen Fällen notwendig, die Wirklichkeit so akkurat wie möglich abzubilden, so dass je nach Komplexität des zu modellierenden Gegenstandes auch die dem Modell zugrundeliegenden Annahmen äußerst kompliziert sein müssen. Der adäquate Grad an Detailgenauigkeit von Modellen hängt somit vom Forschungszweck ab, der mit der Modellierung verfolgt wird (vgl. Axelrod 2007: 93–94; Bonabeau 2002). Morton unterscheidet aus diesem Grund zwischen theoretischen und angewandten Modellen. Theoretische Modelle sind demnach „designed not to be empirically estimated using real-world data but rather to represent a highly stylized version of the real world.“ (Morton 1999: 61) Bei der Entwicklung dieser Modelle wird somit gar nicht der Anspruch erhoben, ein möglichst genaues Abbild der Realität zu erstellen. Angewandte Modelle werden hingegen „designed (a) to provide predictions that can be used as a basis for hypotheses about the real world or (b) to be directly evaluated empirically.“ (Morton 1999: 61)

In der Politikwissenschaft wird zunehmend versucht, Modelle so zu gestalten, dass sie zur Erstellung von Zukunftsprognosen verwendet werden können. Entsprechend werden formale Modelle in der Politikwissenschaft stetig detaillierter in der Nachbildung des jeweiligen Wirklichkeitssegments und ihr Abstraktionsgrad sinkt (vgl. Achen 2006a: 294–295; Hug 2014: 290–291). Daraus folgt jedoch, dass die größere Komplexität der Modellregeln und die häufig wachsende Zahl an berücksichtigten Akteuren die analytische Lösung von *Rational Choice*-Modellen deutlich erschwert, eventuell sogar verunmöglicht. In vielen Fällen lösen Modellierer dieses Problem, indem sie die Anzahl der Akteure reduzieren, oder ähnliche Simplifizierungen durchführen. Dies mag auch für Akteure in der Politik gelten. Negative Konsequenz dieser Vorgehensweise ist jedoch, dass das Modell die Wirklichkeit voraussichtlich inadäquat abbildet und entsprechend die Qualität bzw. Wahrscheinlichkeit der mit dem Modell erstellten Zukunftsprognosen sinkt. Eine Möglichkeit, die Komplexität der Modelle zu erhalten und dennoch Lösungen zu generieren, sind Computersimulationen. Sie dienen dazu, bei derart komplizierten Modellen dennoch ein Gleichgewicht zu ermitteln, ohne dass eine Simplifizierung notwendig ist, die die Analogie von virtueller und wirklicher Welt verringern würde (vgl. Duffy

1992: 241–242, 269; Johnson 1999: 1518; Opp 2015: 207–208). Des Weiteren sind sie ein nützliches Instrument für Forschungsdesigns, die eine Vielzahl an formalen Modellen sowie einen umfangreichen Datensatz umfassen und aus diesem Grund einen großen Aufwand an zeitlichen und personellen Ressourcen benötigen würden, um die Modelle analytisch zu lösen.

Dies gilt auch für das DEU-Projekt, in dem verschiedene formale Modelle für Entscheidungsprozesse in der EU auf einen Datensatz mit 162 Fällen mittels Computerprogrammierungen angewandt worden sind. In Abgrenzung zum vorangegangenen Unterkapitel liegt der Schwerpunkt des DEU-Projekts jedoch nicht auf der Entwicklung von Zukunftsprognosen. Vielmehr soll anhand der Modellierung bereits vergangener Ereignisse zunächst überprüft werden, welche Modelle die Wirklichkeit am geeignetsten abbilden. Daraus werden schließlich grundlegende Charakteristika von EU-Entscheidungsprozessen abgeleitet. Im Folgenden werden das Forschungsdesign sowie die für die vorliegende Arbeit relevanten Forschungsergebnisse des DEU-Projekts erläutert.

2.2 *Rational Choice* Institutionalismus in EU-Studien

Bis zum Beginn der 1990er Jahre lag der Fokus von EU-Studien auf deskriptiven Fallstudien (vgl. Mattila 2012: 453; Pahre 2005: 114). Diese haben zweifellos einen großen Beitrag zum Erkenntnisgewinn über die EU sowie für die empirische Grundlage der EU-Studien geleistet. Dennoch weist dieser methodische Ansatz Schwierigkeiten und Grenzen auf: Fallanalysen beschränken sich häufig auf kurze Phasen politischer Prozesse, die von großen Konflikten oder besonderen Ereignissen in dem jeweiligen Politikfeld geprägt sind. Zudem sind sie in der Regel durch einen hohen Grad an Detailgenauigkeit gekennzeichnet, was es häufig erschwert zu ermitteln, welche Befunde für einen bestimmten Fall einzigartig sind und welche für eine gesamte Klasse von Fällen gelten. Daher besitzen generalisierende Schlussfolgerungen, die auf der Grundlage von Fallstudien gezogen werden, nur begrenzte Validität (vgl. Achen 2006a: 265; Achen/Snidal 1989: 146; Hagemann 2015: 137).

Vor dem Hintergrund dieser Kritik entwickelte sich in den 1990er Jahren ein Forschungszweig, der formale Modelle zur Analyse von europäischen Entscheidungsprozessen anwendet. Ziel dieses Ansatzes ist es, mittels verschiedener formaler Modelle und statistischer Methoden generelle Theorien über Entscheidungsprozesse in der EU zu entwickeln und zu testen (vgl. Stokman/Thomson 2004b: 5–6). Zen-

trale Meilensteine dieses Forschungszweigs sind die Studien „The European Union Decides“ (Thomson et al. 2006b, DEUI) sowie die darauf aufbauende Monographie „Resolving Controversy in the European Union, Legislative Decision-Making Before and After Enlargement“ (Thomson 2011).³¹ Sowohl hinsichtlich der Modellanzahl und der damit verbundenen Spannbreite an theoretischen Erklärungsmustern als auch hinsichtlich des empirischen Datensatzes stellen diese Arbeiten bislang das umfangreichste Projekt dar, in dem formale Modelle zur Analyse von Entscheidungsprozessen in der EU angewandt und verglichen wurden. Entsprechend beurteilt Bueno de Mesquita die DEUI-Studie als „the finest work done thus far in applying rigorous standards to the empirical evaluation of competing explanations of decision-making“ (Bueno de Mesquita 2004: 125; siehe auch Crombez/Vangerven 2014: 297; Mattila 2012; Schneider et al. 2006: 300). Die DEU-Studien geben Aufschluss darüber, welche Modelle in der Nachbildung von EU-Entscheidungsprozessen eine hohe Prognosefähigkeit aufweisen und sich daher für zukünftige Analysen besonders anbieten. Des Weiteren stellen sie einen Bezugspunkt für die Operationalisierung der Inputvariablen von Modellen dar. Die vorliegende Untersuchung baut auf den Erkenntnissen der DEU-Studien auf, indem sie ihre Modellauswahl sowie die methodische Vorgehensweise bei der Operationalisierung der Inputvariablen auf deren Forschungsergebnisse stützt. Im Folgenden werden daher das Forschungsdesign sowie die zentralen Forschungsergebnisse des DEU-Projekts erläutert.

31 Vorläufer der DEUI-Studie ist die Publikation *European Community Decision Making: Models, Applications and Comparisons*, herausgegeben von Bueno de Mesquita und Stokman (1994a), die sich durch ein ähnliches Forschungsdesign auszeichnet. Bueno de Mesquita und Stokman analysieren darin 16 bereits abgeschlossene Entscheidungsprozesse im Rat der EU mittels zweier Verhandlungsmodelle und vergleichen die Modelle anschließend hinsichtlich der Richtigkeit und Präzision der von den Modellen ex post generierten Vorhersagen. Das Ziel dieses Vergleichs besteht darin, zu ermitteln, welches Modell Entscheidungsprozesse im Rat adäquater abbildet und daraus Rückschlüsse auf EU-Entscheidungsprozesse zu ziehen, sowie Modelle zu entwickeln, die für Zukunftsprognosen in der EU-Politik verwendet werden können. Thomson et al. (2006) lehnen ihre Studie an dieses Forschungsdesign an, weiten die Forschung aber aus, indem sie eine größere Anzahl an Modellen auf ein wesentlich größeres Datenset testen (vgl. Achen 2006a: 295; Thomson/Hosli 2006: 8–9 sowie Abschnitt 1.2.2).

Forschungsdesign

Untersuchungsgegenstand der Studien von Thomson et al. (2006) und Thomson (2011) sind tägliche Politikentscheidungen in der EU, die bis dahin ein Forschungsdesiderat ausmachten, nach Ansicht der Herausgeber aber zentral für das Verständnis der Entwicklung der EU als supranationale Institution sind:

[... U]nderstanding the quotidian political battles of the EU is not of minor concern. To the contrary, the relentless march of daily decision-making has at least as much importance and deserves every bit as much attention as the sporadic grand bargains that dot recent European history. Both are critical to understanding why the EU has had the impact it has had, and why it is among the world's most effective supranational political organisations. But only the grand bargains have enjoyed extensive discussion. Hence the focus of this book: how can we explain everyday decisions in the European Union? (Schneider et al. 2006: 300)

Die Analyse erfolgt durch die Anwendung von formalen Modellen. Eine besondere Qualität von formalen Modellen gegenüber anderen methodischen Ansätzen besteht zum einen in ihrer Präzision, da Annahmen über den politischen Entscheidungsprozess widerspruchsfrei in mathematischen Formeln ausgedrückt werden müssen. Dies verhindert zudem, dass die der Forschung zugrundeliegenden Annahmen zwar in der Analyse impliziert, aber nicht explizit benannt werden:

When researchers set forth their ideas about EU decision-making as models, they must be explicit about the propositions they make regarding the decision-making processes at work, as well as the assumptions contained in their models. Assumptions in verbally formulated theories are often implicit rather than absent. Modelling allows the implications of the propositions to be drawn out through deductive reasoning. (Thomson/Hosli 2006: 10; für eine ähnliche Argumentation siehe u.a. Kroneberg/Kalter 2012)

Des Weiteren erfordert die Darstellung des Entscheidungsprozesses in der Sprache formaler Modellierungen notwendigerweise eine Reduzierung von Komplexität (vgl. die Ausführungen in Abschnitt 2.1). Dies mag als Beschränkung erscheinen, birgt aber zugleich Vorteile: Der Modellierer wird dazu angehalten, sich in der Analyse auf die wesentlichen Merkmale des Entscheidungsprozesses zu konzentrieren.

ren und Variablen, die für den Einzelfall interessant sein mögen, die Analyse der grundlegenden Mechanismen des Prozesses aber verschleiern, unberücksichtigt zu lassen (vgl. Pahre 2005: 120).³²

Die formalen Modelle, die in der DEUI-Studie berücksichtigt werden, beinhalten divergierende Annahmen über die wesentlichen Bestandteile von EU-Entscheidungsprozessen (vgl. Thomson/Hosli 2006: 9). Sie alle werden auf den gleichen Datensatz angewandt, der aus 162 kontroversen Themen verhandelt im Rahmen von 66 Gesetzesinitiativen aus den Jahren 1999 bis 2001 besteht (vgl. Thomson et al. 2006a: xvii; Thomson/Hosli 2006: 12–13).³³ Die Modelle werden dazu eingesetzt, den Entscheidungsprozess nachzubilden und ex post die jeweiligen Verhandlungsergebnisse „vorherzusagen“. Anschließend werden die Modelle hinsichtlich der Genauigkeit ihrer Vorhersagen mittels statistischer Methoden verglichen.³⁴ Aus der Genauigkeit der Vorhersagen kann anschließend abgeleitet werden, welches Modell den Entscheidungsprozess am adäquatesten abbildet. Auf diese Weise können Hypothesen über generelle Charakteristika von EU-Entscheidungsprozessen getestet und Regelmäßigkeiten identifiziert werden (vgl. Thomson/Hosli 2006: 9–11). Insgesamt verbindet der Forschungsansatz somit Fallstudien als Inputdaten, formale Modelle für die Analyse der Entscheidungsprozesse sowie statistische Methoden zum Vergleich ihrer Prognosefähigkeit (vgl. Achen 2006a: 264). Das Forschungsdesign steht daher – auch wenn die zuvor genannte Kritik dies fälschlicherweise suggerieren mag – nicht im Gegensatz zu deskriptiven Studien über EU-Entscheidungsprozesse. Vielmehr baut es auf den Erkenntnissen jener Studien auf, da diese sowohl für die Entwicklung als auch die Verbesserung der Modelle von großer Bedeutung sind. Dennoch liefert der

32 Siehe dazu auch Crombez und Vangerven (2014), die einen Überblick über die EU-Forschung mittels verfahrensrechtlicher Modelle in den vergangenen zwei Jahrzehnten liefern und darin aufzeigen, inwiefern formale Modelle dazu beitragen konnten, die grundlegenden Mechanismen der in den EU-Verträgen festgeschriebenen Verfahrensweisen in Entscheidungsprozessen präzise herauszuarbeiten.

33 Für eine kritische Diskussion der Fallauswahl siehe u.a. Hagemann (2015) und Princen (2012).

34 Für eine ausführliche Erklärung zur Vorgehensweise dieser Messung siehe Achen (2006a: 293) sowie Thomson (2011: 179).

beschriebene Ansatz einen zusätzlichen Erkenntnisgewinn, indem er die Möglichkeit schafft, generalisierende Schlussfolgerungen über EU-Entscheidungsprozesse zu ziehen (vgl. Thomson/Hosli 2006: 10–11)³⁵.

Die DEUII-Studie folgt dem gleichen Forschungsansatz. Sie divergiert aber insofern von der DEUI-Studie, als sie eine geringere Anzahl an Modellen sowie eine größere Anzahl an Fällen einbezieht. Während das DEUI-Projekt lediglich Entscheidungsprozesse der EU-15-Periode analysiert, berücksichtigt der Datensatz des DEUII-Projektes auch Entscheidungsprozesse der EU-25- und EU-27-Perioden.³⁶ Dies ermöglicht es zu untersuchen, ob die Erweiterung der EU Veränderungen hinsichtlich der Entscheidungsprozesse bewirkt hat (vgl. Thomson 2011: 160).³⁷

Die theoretischen Grundlagen beider DEU-Studien sowie der darin berücksichtigten Modelle bildet der *Rational Choice* Institutionalismus. Die mit diesem Ansatz verbundenen Grundannahmen werden im Folgenden dargelegt. Daran anschließend werden die verschiedenen Kategorien von *Rational Choice*-Modellen, die das DEU-Projekt einbezieht, erläutert.

Rational Choice Institutionalismus

Der *New Institutionalism* hat sich am Ende der 1970er sowie in den 1980er Jahren aus einer Kritik an den vorherrschenden behavioristischen Perspektiven entwickelt und zielt darauf ab, den Einfluss von Institutionen auf soziale und politische Phänomene zu analysieren. Er gliedert sich in drei Ansätze auf, die auf äußerst unterschiedlichen Annahmen basieren und verschiedene Schwerpunkte in ihren Analysen setzen: der Historische Institutionalismus, der Soziologische Institutionalismus und der *Rational*

35 Um die grundlegenden Charakteristika des Forschungszweigs zu erklären, an den die vorliegende Arbeit anknüpft, wurde der Schwerpunkt auf die DEU-Studien gelegt. Als weitere umfassende Publikation ist aus dem DEU-Projekt ein Special Issue im Journal *European Union Politics* mit dem Titel „Winners and Losers in the European Union“ hervorgegangen, das Erkenntnisse auf der Akteursebene bereitstellt. Es beschäftigt sich u.a. mit Fragen zu den Machtverhältnissen zwischen den relevanten Akteuren, zu ihrem Einfluss auf die Verhandlungsergebnisse etc. (vgl. Stokman/Thomson 2004b: 5–6). Darüber hinaus, wird auf den Forschungsansatz im Allgemeinen sowie den Datensatz im Besonderen in verschiedenen weiteren Veröffentlichungen Bezug genommen, die in Teilen in der Diskussion der Inputvariablen (vgl. Abschnitte 4.3 bis 4.6) aufgegriffen werden.

36 Für eine ausführliche Beschreibung des DEUII-Datensatzes siehe auch Thomson et al. (2012).

37 Für einen Vergleich von Forschungsdesign und -ergebnissen der DEUI- und DEUII-Studie siehe Mattila (2012).

Choice Institutionalismus (im Folgenden: RCI) (vgl. Hall/Taylor 1996: 936). Das DEU-Projekt sowie die vorliegende Arbeit sind letzterem Ansatz zuzuordnen, der im Folgenden skizziert wird.

Der RCI richtet sich auf die Beantwortung von drei Forschungsfragen: Welchen Einfluss haben Institutionen auf das Handeln von Akteuren? Warum sind Institutionen notwendig? Warum entstehen Institutionen und warum bestehen sie weiter? In Anlehnung an diese Forschungsfragen lassen sich zwei Analyseebenen des RCI unterscheiden: Die erste Strömung sieht Institutionen als exogen gegeben an und analysiert ihre Auswirkungen. Dies ist der dominante Ansatz, der bereits sehr weit entwickelt ist und an den die vorliegende Arbeit anknüpft. Die zweite Strömung untersucht Institutionen als endogene Variable und fragt nach den Ursachen ihrer Entstehung, ihrer konkreten Ausgestaltung sowie ihres weiteren Bestandes (vgl. Shepsle 2006: 24–25; Weingast 2002: 661, 670, 691).

Beide Strömungen gründen auf bestimmte Verhaltensannahmen von Individuen: Es wird angenommen, dass die Akteure über Präferenzen bezüglich der ihnen zur Disposition stehenden Handlungsalternativen verfügen und zu diesen Präferenzen konforme politische Forderungen stellen. Sie verfolgen das Ziel, ihren Nutzen entsprechend ihrer Präferenzen zu maximieren. Dabei handeln sie strategisch, d.h. sie berücksichtigen bei ihren Entscheidungen, welches Handeln von den anderen Akteuren zu erwarten ist und wie ihr eigenes Handeln das Handeln der anderen Akteure beeinflussen könnte. Institutionen bilden den Rahmen der Möglichkeiten, in dem sich die Akteure bewegen können und beeinflussen daher ihr Handeln. Politikergebnisse sind somit durch das Zusammenspiel von Akteursinteressen und Institutionen bestimmt (vgl. Dehling/Schubert 2011: 31; Hall/Taylor 1996: 944–945; Mühlböck/Rittberger 2015: 6; Thomson 2011: 14; Thomson/Hosli 2006: 9; Weingast 2002: 661–662).

In den RCI-Publikationen herrscht keine Einigkeit hinsichtlich der Definition von Institutionen. Die zentrale Konfliktlinie verläuft in dieser Debatte zwischen Wissenschaftlern, die Institutionen als organisatorische Einheit betrachten, wie z. B. den Europäischen Rat, den Bundestag oder politische Parteien, sowie Wissenschaftlern, die unter dem Begriff der Institution auch informelle Strukturbestandteile wie Regeln und Normen fassen (vgl. Mühlböck/Rittberger 2015: 6; Ostrom 2007: 23). Dementsprechend differenziert Shepsle zwischen strukturierten und unstrukturierten Institutionen (vgl. Shepsle 2006: 27). Eine sehr weite – aber einschlägige

– Bestimmung von Institutionen liefert North, der in seiner Definition beide Pole zusammenführt: „Institutions are the humanly devised constraints that structure political, economic and social interaction. They consist of both informal constraints (sanctions, laws, customs, traditions, and codes of conduct), and formal rules (constitutions, laws, property rights).“ (North 1991: 97)

In den Modellen des DEU-Projekts werden beide Gruppen von Institutionen modelliert. Diese Arbeit folgt daher der umfassenden Definition von North. Gleichwohl ist anhand der Modellkategorien in der Regel zu erkennen, welcher Schwerpunkt jeweils bei der Modellierung von Institutionen – ob formell oder informell – gesetzt wird. So lassen sich drei Modellkategorien voneinander abgrenzen, deren maßgebliches Unterscheidungskriterium in der Berücksichtigung bzw. Nicht-Berücksichtigung formaler Gesetzgebungsprozesse besteht. Diese Modellkategorien werden im Folgenden präsentiert.

Modellkategorien

Die für das DEU-Projekt ausgewählten Modelle sollen dazu dienen, den Entscheidungsprozess im Rat der EU nachzubilden. Sie alle basieren auf dem zuvor beschriebenen *Rational Choice*-Ansatz, der in der Politikwissenschaft zunehmend angewandt wird und in den letzten Jahren einen großen Beitrag zum Erkenntnisgewinn in EU-Studien geleistet hat. Dies äußert sich u.a. darin, dass sich *Rational Choice*-Modelle in der Regel durch eine hohe Vorhersagekraft auszeichnen (vgl. Schneider et al. 2011: 7, 2006: 300; Veen 2011a: 22).

Trotz ihrer identischen theoretischen Fundierung repräsentieren die Modelle unterschiedliche Annahmen über den Prozess, durch den Politikpräferenzen in politische Entscheidungen und Gesetze umgewandelt werden (vgl. Thomson/Hosli 2006: 6). Es wird davon ausgegangen, dass der Entscheidungsprozess aus zwei Phasen besteht: Die erste Phase umfasst informelle Verhandlungen, in denen die Spieler Informationen sammeln und austauschen, ihren Verhandlungspartnern drohen oder Kompromisse aushandeln können etc. Sofern die Verhandlungspartner sich auf ein vorläufiges Ergebnis einigen, beginnt die zweite Phase. Diese wird durch den formalen Gesetzgebungsprozess gestaltet. Die beiden Phasen stehen jedoch nicht vollständig abgeschlossen nebeneinander, sondern beeinflussen sich wechselseitig, da die formalen Abstimmungsregeln der zweiten Phase die beteiligten Akteure sowie deren Machtverhältnisse in der ersten Phase bedingen (vgl. Achen 2006b: 86–87). Viele Modelle betrachten die beiden Phasen dennoch getrennt voneinan-

der, weshalb sich drei Modellkategorien unterscheiden lassen: *procedural models* (im Folgenden: verfahrensrechtliche Modelle), *bargaining models* (im Folgenden: Verhandlungsmodelle) und *mixed models* (im Folgenden: Hybridmodelle).

- *Verfahrensrechtliche Modelle*: Verfahrensrechtliche Modelle legen bei der Entwicklung der Modellregeln den Schwerpunkt auf die konkreten rechtlichen Abläufe des Gesetzgebungsprozesses. Das bedeutet, dass in diesen Modellen die Spielerauswahl, ihre Rechte im Entscheidungsprozess sowie die Sequenz der Bestandteile des Gesetzgebungsprozesses genau an den formalen Gesetzesregeln angelehnt sind (vgl. Crombez/Vangerven 2014: 290, 294; Schneider et al. 2006: 300; Steunenber/Selck 2006: 80).
- *Verhandlungsmodelle*: Verhandlungsmodelle legen bei der Entwicklung der Modellregeln den Schwerpunkt auf die informellen Verhandlungen, die vor dem formalen Gesetzgebungsprozess stattfinden. Dennoch berücksichtigen derlei Modelle, dass die formalen Regeln gewisse Grenzen für die informellen Verhandlungen setzen. So spielen sie zwar für die Sequenz der Modellregeln keine Rolle, bestimmen aber, welche Akteure in die Analyse mit einbezogen werden und wie viel Gewicht ihren Positionen zugeschrieben wird. In den Verhandlungsmodellen des DEU-Projekts werden ausschließlich Akteure berücksichtigt, die auch im formalen Prozess Stimmrechte besitzen. Zwar mögen beispielsweise bestimmte Interessengruppen die Positionen der Entscheidungsakteure beeinflusst haben, aber es wird angenommen, dass diese Form der Einflussnahme vor den informellen Verhandlungen stattfindet, die in diesen Modellen nachgebildet werden (vgl. Schneider et al. 2006: 301; Veen 2011a: 23–24).
- *Hybridmodelle*: Hybridmodelle verbinden verfahrensrechtliche und Verhandlungsmodelle. Sie berücksichtigen in ihren Modellregeln, dass informelle Verhandlungen den Entscheidungsprozess beeinflussen, die abschließende Lösung verbleibender Kontroversen sowie die Festlegung des tatsächlichen Verhandlungsergebnisses aber erst im formalen Gesetzgebungsprozess erfolgt (vgl. Schneider et al. 2006: 301).

Die Modelle lassen sich des Weiteren danach unterscheiden, ob sie der kooperativen oder der nicht-kooperativen Spieltheorie zuzuordnen sind. In kooperativen Spielen können Spieler verbindliche Absprachen treffen, da exogene Mechanismen

bestehen, mittels derer die Einhaltung von Verträgen bindend durchgesetzt werden kann. In der politischen Sphäre handelt es sich dabei meist um Rechtssysteme, im Rahmen derer die entsprechenden Institutionen Vertragsverletzungen überprüfen und gegebenenfalls Sanktionen verhängen können. In der kooperativen Spieltheorie wird angenommen, dass Verträge aufgrund der Höhe der Sanktionen in jedem Fall eingehalten werden. In nicht-kooperativen Spielen sind hingegen keine bindenden Vereinbarungen möglich. Das bedeutet nicht, dass Kooperationen in nicht-kooperativen Spielen nicht möglich sind. Sie entstehen in der nicht-kooperativen Spieltheorie aber nicht als exogen vorgegebene Spielregel zur unbedingten Einhaltung von Verträgen, sondern lediglich als Resultat von Entscheidungen der Spieler (vgl. Holler/Illing 2009: 22; Rieck 2012: 35–37; Sieg 2010: 91).

Alle Modelle sind durch eine räumliche Vorstellung von Politik gekennzeichnet.³⁸ Diese existiert schon lange, ihre bekannteste politikwissenschaftliche Systematisierung erfuhr sie jedoch erstmals durch Downs' „An Economic Theory of Democracy“ (1957) sowie Blacks „The Theory of Committees and Elections“ (1958). Besonders geläufig ist gegenwärtig – auch jenseits der Politikwissenschaft – die räumliche Vorstellung von Parteipolitik auf einem ideologischen Links-Rechts-Kontinuum. Beim DEU-Projekt handelt es sich um die einfachste Form räumlicher Modellierung, nämlich um *eindimensionale* räumliche Modelle. Der modellierte Politikraum entspricht dem jeweiligen kontroversen Thema, das in den Entscheidungsprozessen verhandelt wird. Die Gesamtheit der Policypositionen, die bezüglich jener politischen Sachfragen eingenommen werden können, werden auf einer numerischen Skala dargestellt, die im DEU-Projekt auch als „issue continuum“ (Thomson 2011: 22) bezeichnet wird. Die Distanz zwischen den Policypositionen in diesem Kontinuum muss dabei die tatsächliche „politische Entfernung“ zwischen den entsprechenden Politikalternativen reflektieren. Sofern sich die Policypositionen direkt mit Zahlen identifizieren lassen – wie z. B. bei der Festsetzung einer Steuer – ist die Unterteilung des Zahlenraums *a priori* gegeben; werden auf der Skala hingegen abstrakte Politikdimensionen repräsentiert, sind Interpretationen durch den Modellierer not-

38 Eine Einführung in die Grundlagen der räumlichen Modellierung bieten Linhart, Kittel und Bächtiger (2014).

wendig.³⁹ Die Policypositionen der Entscheidungsakteure werden anschließend in das Kontinuum eingetragen. Es wird vorausgesetzt, dass die Präferenzordnungen bzw. Nutzenfunktionen der Spieler vollständig und transitiv sind. Des Weiteren wird angenommen, dass die Akteure eingipflige Präferenzen haben. Dies impliziert, dass die Akteure ausschließlich einen Idealpunkt besitzen und den größten Nutzen durch dasjenige Politikergebnis erhalten, das ebendiesem Idealpunkt entspricht. Demgegenüber werden Politikergebnisse umso weniger von den Akteuren bevorzugt, je weiter sie in dem Kontinuum von ihrer eigenen Position entfernt sind, d.h. mit zunehmender Distanz vom Idealpunkt nimmt der Nutzen entsprechend ab. (vgl. Behnke 2009: 189; Linhart 2014: 14–15; Thomson 2011: 21–23, 40) Eine räumliche Konzeptionalisierung von politischen Sachfragen ist wichtig, um die Prognosefähigkeit von Modellen vergleichen zu können, da auf diese Weise überprüft werden kann, ob das prognostizierte Politikergebnis dem tatsächlichen Politikergebnis entspricht, bzw. wie weit – in den Kategorien der räumlichen Modellierung – die Prognose von der Wirklichkeit abweicht (vgl. Thomson/Hosli 2006: 25).

Nachdem nun die theoretischen Grundlagen des Projekts skizziert worden sind, wird im Folgenden die Erhebung der empirischen Daten des Projekts erläutert.

Inputdaten

Die Inputdaten für die formalen Modelle basieren im DEU-Projekt auf qualitativen Erhebungen. Für die DEUI-Studie wurden Daten zu 66 Gesetzesinitiativen erhoben, die zwischen 1999 und 2001 verhandelt wurden und insgesamt 162 kontroverse Sachfragen thematisieren. Die Phase der Entscheidungsprozesse, die bei der Datenerhebung berücksichtigt wurde, begann mit der Vorlage des Gesetzesvorschlags durch die Europäische Kommission und endete mit der finalen Entscheidung über den Vorschlag, d.h. mit Annahme oder Ablehnung durch den Rat und/oder das Europäische Parlament. Die Gesetzesvorschläge beziehen sich auf verschiedene Poli-

39 Bei der Gestaltung räumlicher Modelle ist eine konsistente Einordnung der Politikoptionen in dem Kontinuum notwendig vorausgesetzt. Zur Veranschaulichung dieser Schwierigkeit führt Linhart beispielhaft verschiedene Fälle vor, in denen eindimensionale räumliche Modellierungen angemessen bzw. nicht möglich oder zumindest nicht sinnvoll sind. Aus diesen Erläuterungen lassen sich für Wissenschaftler, die an der Entwicklung von räumlichen Modellen interessiert sind, hilfreiche Hinweise für die konkrete Definition des Kontinuums ableiten (vgl. Linhart 2014).

tikfelder (vgl. Thomson/Hosli 2006: 12–13). Die Quantität an analysierten Fällen sowie deren inhaltliche Spannbreite sind für den Forschungsansatz von zentraler Bedeutung, um die Prognosefähigkeit der Modelle zu vergleichen:

The aim of this study is to apply and compare different explanations of legislative decision-making in the European Union (EU). Two features of the research design are particularly important with respect to achieving this aim. First, the selection of cases must cover a sufficient number and variety of cases to count as test of the explanations. Second, a way of thinking about very different decision situations has to be devised, such that they can be compared, in terms of the applicability of different explanations in any given situation, and in terms of the performance of explanations in different situations. (Thomson/Stokman 2006: 25)

Weitere Kriterien für die Auswahl der Fälle bestanden in der Kategorie des Gesetzgebungsprozesses, dem Zeitpunkt, in dem das Gesetz vorgeschlagen wurde und ihrer politischen Bedeutung sowie dem Grad der Kontroverse. Die Erhebung der Daten erfolgte mittels 150 halbstrukturierter Interviews mit 125 Experten aus den EU-Mitgliedstaaten, der Kommission, dem Europäischen Parlament und dem Generalsekretariat des Rates der Europäischen Union. Die von den Experten bereitgestellten Informationen betrafen die Policypositionen der Entscheidungsakteure zu Beginn der jeweiligen Verhandlungen, die Priorität, die die Akteure den Themen verliehen haben sowie die Fähigkeiten bzw. Ressourcen der Akteure, um ihre Positionen durchzusetzen. Die Fähigkeiten der Akteure wurden zudem mittels zweier Varianten des Shapley Shubik Index operationalisiert und deren Adäquatheit zur Widerspiegelung der Akteursfähigkeiten im Rahmen des Modellvergleichs ebenfalls gegenübergestellt (vgl. Thomson/Stokman 2006: 31–48).

Thomson (2011) verwendet in der DEUII-Studie sowohl Daten aus dem DEUI-Projekt als auch einen zusätzlich erhobenen Datensatz. Insgesamt wurden 125 Gesetzesvorschläge mit 331 Themen für die DEUII-Studie ausgewählt, von denen 69 aus der EU-15-Periode und 56 aus der EU-25- bzw. EU-27-Periode stammen. Die Daten wurden ebenfalls durch Experteninterviews erhoben (vgl. Thomson 2011: 27).⁴⁰ Im Folgenden werden die für die vorliegende Arbeit relevanten Forschungsergebnisse des DEU-Projekts erläutert.

40 Für eine ausführliche Erläuterung von methodischen Unterschieden bei der Datenerhebung zwischen der DEUI- und der DEUII-Studie siehe Mattila (2012: 456).

Zentrale Forschungsergebnisse

Das DEU-Projekt hat wichtige Erkenntnisse über Entscheidungsprozesse in der EU generieren sowie bereits entwickelte Hypothesen aus vorherigen Studien bestätigen bzw. widerlegen können. An dieser Stelle wird aufgrund des Umfangs der Forschungsergebnisse auf eine vollständige Darstellung ebendieser verzichtet. Stattdessen wird der Schwerpunkt auf Erkenntnisse gelegt, die für den weiteren Verlauf der vorliegenden Untersuchung relevant sind. Diese umfassen zum einen Forschungsergebnisse zur Prognosefähigkeit der verschiedenen Modellkategorien, da sie Hinweise bezüglich der Frage liefern, welche Modelle dazu geeignet sein könnten, die Erdgasaußenpolitik der EU abzubilden. Dies gilt speziell für die Untergliederung von Verhandlungsmodellen, die aufgrund ihrer Abbildung von informellen Verhandlungsprozessen für die intergouvernementale Erdgasaußenpolitik der EU besonders geeignet sind. Zum anderen werden Forschungsergebnisse erläutert, die die Akteursauswahl der vorliegenden Arbeit betreffen.

Durch den Vergleich der Prognosefähigkeit verschiedener Modelle konnte festgestellt werden, dass Verhandlungsmodelle als Modellklasse insgesamt die akkuratesten Prognosen liefern. Die Prognosen von verfahrensrechtlichen Modellen waren hingegen wesentlich ungenauer: „[...A]t this stage of our understanding, pure bargaining models predict best as a class, while pure legalist models perform worst. The more mixed models include procedural aspects, the worse they do.“ (Schneider et al. 2006: 303–304) Dies gilt auch für die zusätzlichen Fälle der DEUII-Studie, in denen die Prognosefähigkeit der Verhandlungsmodelle konstant bleibt. Daraus wurde geschlossen, dass der Entscheidungsprozess in der EU durch formale Regeln allein

nicht adäquat widergespiegelt wird, sondern informelle Regeln und Verfahrensweisen – sowohl vor als auch nach der Erweiterung – von großer Bedeutung sind (vgl. Achen 2006a: 295; Thomson 2011: 185).⁴¹

Bezüglich der Prognosefähigkeit von Verhandlungsmodellen wurden zwei weitere Spezifika ermittelt: Erstens sind die Prognosen von kooperativen Verhandlungsmodellen insgesamt besser als die von nicht-kooperativen Verhandlungsmodellen. Zweitens ist die Prognosefähigkeit von denjenigen Modellen am höchsten, die als Simulationsergebnis einen Mittelwert – unabhängig davon ob gewichtet oder ungewichtet – generieren. Daraus lässt sich ableiten, dass in EU-Entscheidungsprozessen alle Positionen – auch abweichende bzw. extreme – mit einbezogen und nicht übergangen werden (vgl. Achen 2006a: 297). Informelle Verhandlungen und Kompromissfindung sind somit zentrale Charakteristika der Entscheidungsfindung – auch in Fällen, in denen einstimmige Entscheidungen formalrechtlich nicht erforderlich sind: „*Unanimity, wherever possible, is a very strong norm in the EU, even when decision outcomes supported by only a qualified majority of actors are possible.*“ (Stokman/Thomson 2004b: 19; siehe auch Bueno de Mesquita 2004: 133) Dies mag vor allem darin begründet sein, dass EU-Akteure aufgrund der institutionellen Verflechtungen in regelmäßigen Verhandlungen auf diversen Politikfeldern stehen und daher wissen, dass sie auch in zukünftigen Entscheidungsprozessen auf Ko-

41 Slapin (2014) zweifelt dieses Ergebnis an. Er argumentiert, dass die Positionen von EU-Kommission und -Parlament höhere Messfehler aufwiesen als die Positionen der Mitgliedstaaten. Verfahrensrechtliche Modelle, die Kommission und Parlament einen großen Einfluss zuschreiben, wiesen aus diesem Grund eine geringere Prognosefähigkeit auf als Verhandlungsmodelle. Leinaweaver und Thomson (2014) weisen diese Kritik zurück. Sie merken aber durchaus an, dass verfahrensrechtliche Modelle nicht *grundsätzlich* schlechter geeignet seien, um EU-Entscheidungsprozesse abzubilden. Vielmehr weisen Verhandlungsmodelle eine hohe Prognosefähigkeit bei besonders kontroversen Themen auf, die den DEU-Datensatz maßgeblich ausmachen, während verfahrensrechtliche Modelle hingegen diejenigen Fälle besser abbilden würden, in denen keine große Kontroverse mehr zwischen den Entscheidungsakteuren bestehe (siehe auch Crombez/Vangerven 2014). Aus dieser Auseinandersetzung lässt sich schließen, dass EU-Entscheidungsprozesse sich sowohl durch formale als auch informelle Aspekte auszeichnen, was auch Hagemann (2015: 148) in einem Forschungsüberblick zu Verhandlungen in der EU festhält: „So rather than a “regime” of *either* informal *or* formal arrangements, Council bargaining is a complex mix of the two modes of governance [...]“. Für die vorliegende Untersuchung und speziell für die Modellauswahl in Abschnitt 4.1 ist diese Debatte allerdings nicht von praktischer Relevanz, da die EU-Erdgasaußenpolitik auf intergouvernementaler Ebene und nicht in einem ordentlichen Gesetzgebungsverfahren verhandelt wird, verfahrensrechtliche Modelle als Abbild dieses Policy-Subsystems also *a priori* ausgeschlossen werden können.

operation und Zugeständnisse von anderen Akteuren für die Durchsetzung ihrer Interessen abhängen: „Norms of trust and cooperation naturally emerge because EU participants know that, in addition to the legalistic procedural constraints, they must interact with one another in the future and therefore stand to gain by learning to cooperate and compromise.“ (Bueno de Mesquita 2004: 133; siehe auch Schneider et al. 2006: 305)

Des Weiteren lassen die erfolgreichen Prognosen bestimmter Verhandlungsmodelle darauf schließen, dass die Macht der Akteure sowie die Priorität, die sie den jeweiligen Verhandlungsthemen verleihen, ebenfalls substanziellen Einfluss darauf haben, welche Akteurspositionen stärker berücksichtigt werden (vgl. Achen 2006a: 297). Da der Einfluss der Akteure im DEU-Projekt mittels Indizes bemessen wurde, die die Abstimmungsmacht der Mitgliedstaaten im Rat als entscheidende Variable einbeziehen, ist daraus zu schließen, dass die verfahrensrechtlichen Regeln weniger hinsichtlich des genauen Ablaufs des Entscheidungsprozesses relevant sind, sondern vielmehr bezüglich der Frage, welche Akteure Einfluss auf den Entscheidungsprozess ausüben und wie groß ebendieser Einfluss in Relation zu dem der anderen Akteure ist (vgl. Stokman/Thomson 2004b: 19).

In Bezug auf die Relevanz der Akteure für die Entscheidungsfindung deuten die Forschungsergebnisse darauf hin, dass die Präferenzen der Mitgliedstaaten zwar noch immer von herausragender Bedeutung für die Verhandlungsergebnisse sind. Allerdings verbessern sich die Prognosen der Modelle in der Regel, sobald auch die Positionen der Kommission sowie des Parlaments Berücksichtigung finden: „A strictly defined intergovernmental perspective is therefore no longer helpful for understanding the day-to-day decision-making in the EU. The preferences of the Commission and Parliament matter. The daily decision-making in the Union can therefore be characterized as a mixture of intergovernmental and supranational bargaining.“ (Schneider et al. 2006: 314) Dennoch betont Thomson in der DEUII-Studie, dass die Positionen der Mitgliedstaaten auch nach der EU-Erweiterung noch immer deutlich gewichtiger sind als die Positionen von Kommission und Parlament (vgl. Thomson 2011: 281–282).

Die Präsentation der zentralen Forschungsergebnisse des DEU-Projekts verdeutlicht, wie wirkungsvoll der Forschungsansatz dafür eingesetzt werden kann, um generalisierende Schlussfolgerungen über die grundlegenden Mechanismen des EU-Entscheidungsprozesses zu ziehen. Für die vorliegende Untersuchung liefert das

Projekt auf diese Weise drei wichtige Erkenntnisse: Erstens zeigt es auf, dass von den getesteten Modellen im DEU-Projekt kooperative Verhandlungsmodelle mit einem Mittelwert als Simulationsergebnis im Vergleich zu anderen Modellen die höchste Prognosefähigkeit aufweisen. Dieses Ergebnis muss bei der Diskussion der Modellauswahl berücksichtigt werden. Zweitens weist es darauf hin, dass Einstimmigkeit und Kompromissfindung zentrale Charakteristika von EU-Entscheidungsprozessen sind, was für die Bestimmung der Kompromissbereitschaft der Akteure relevant sein wird. Drittens wurde festgestellt, dass EU-Mitgliedstaaten die wichtigsten Akteure im Entscheidungsprozess darstellen, wobei EU-Kommission und EU-Parlament in ordentlichen Gesetzgebungsverfahren aber nicht unberücksichtigt bleiben sollten. Dies gilt es bei der Akteursauswahl in dieser Untersuchung sowie bei der Gewichtung ihres Einflusses zu berücksichtigen. Gleichzeitig muss aber auch betont werden, dass im weiteren Untersuchungsverlauf ein kritischer Umgang mit den erläuterten Forschungsergebnissen hinsichtlich ihrer Übertragung auf das Policy-Subsystem Erdgasaußenpolitik notwendig ist. Dies ist zum einen darin begründet, dass es sich beim DEU-Projekt um die Analyse von ordentlichen Gesetzgebungsverfahren handelt, bei der Erdgasaußenpolitik der EU hingegen um Entscheidungsprozesse auf intergouvernementaler Ebene. Sich daraus möglicherweise ergebende Differenzen müssen bei der Entwicklung des *Gas Game* bedacht werden. Des Weiteren muss betont werden, dass trotz der an verschiedenen Stellen betonten Bedeutung des DEU-Projekts die Prognose von politischen Entwicklungen bislang insgesamt noch sehr ungenau ist. Die Mehrzahl der im DEU-Projekt getesteten Modelle weist in ihren Prognosen eine höhere Fehlerrate auf als die *baseline* Modelle bestehend aus Mittelwert bzw. Median. So generiert kein Modell substantiell bessere Prognosen als der Mittelwert (vgl. Achen 2006a: 277, 295). Achen betont jedoch, dass diese Schwierigkeiten beim Erstellen präziser Prognosen nicht nur auf formale Modelle, sondern auch auf andere methodische Ansätze zutreffen:

[..]social scientists are very far from predicting political decisions accurately. [...] Nor is that finding solely a judgement on formal models. Neither case studies nor statistical modelling have pointed the way to better predictions. In EU studies as in political science as a whole, we are far from having the conceptual tools of any methodological type that we need to forecast political decision-making well. (Achen 2006a: 295)

Dies soll nicht den Erkenntniswert des DEU-Projekts infrage stellen, sondern darauf hinweisen, dass weitere Forschung in diesem Bereich notwendig ist, was trotz der zahlreichen Publikationen, die an das DEU-Projekt angeknüpft haben, aufgrund der Vernachlässigung von Zukunftsszenarien in ebendiesen weiterhin gilt (vgl. Abschnitt 1.2.2).

In dieser Arbeit wird untersucht, ob das PG, eine spieltheoretische Simulation, die in anderen Kontexten internationaler Verhandlungen bereits eine hohe Prognosefähigkeit bewiesen hat, für die Analyse von EU-Entscheidungsprozessen geeignet ist und somit einen Beitrag zum weiteren Erkenntnisgewinn in diesem Forschungsbereich leisten kann. Dazu werden zunächst die Grundannahmen des Modells, seine formale Herleitung sowie die Ergebnisse seiner Anwendung auf den DEU-Datensatz erläutert.

2.3 Das *Predictioneer's Game*

In der vorliegenden Arbeit soll ein Modell der europäischen Erdgasaußenpolitik gegenüber Russland erstellt werden, um Rückschlüsse aus dem Modell auf den wirklichen Entscheidungsprozess zu ermöglichen und ein Zukunftsszenario zur Entwicklung der europäischen Erdgasaußenpolitik infolge der Ukraine-Krise zu entwickeln. Im vorangegangenen Unterkapitel wurde gezeigt, welche Modelle im Test auf den DEUI-Datensatz die höchste Prognosefähigkeit aufgewiesen haben. Das DEU-Projekt liefert auf diese Weise einen Bezugspunkt für darauf aufbauende Analysen von EU-Entscheidungsprozessen mittels formaler Modellbildung. Es wurde allerdings auf zwei Schwierigkeiten hingewiesen, die sich für die Übertragung der Forschungsergebnisse auf die vorliegende Untersuchung ergeben: Zum einen handelt es sich bei den Modellen mit der höchsten Prognosefähigkeit in erster Linie um *kooperative* Verhandlungsmodelle, die dementsprechend gut dazu geeignet sind, den kooperativen Charakter in EU-Entscheidungsprozessen im Rat widerzuspiegeln. Die Erdgasaußenpolitik der EU ist demgegenüber allerdings – wie in Kapitel 3 noch ausführlich dargelegt wird – durch große Konflikte zwischen den EU-Mitgliedstaaten geprägt, die dazu beigetragen haben, dass sich die EU-Mitgliedstaaten in diesem Policy-Subsystem bislang auf keine kohärente Politik einigen konnten und deswegen nicht dazu bereit waren, Souveränität in großem Maßstab auf die supranationale Ebene zu übertragen. Des Weiteren ist die Prognosefähigkeit von keinem der im DEU-Projekt getesteten Modelle substantiell höher als die des Mittelwerts. Vor

diesem Hintergrund wird im Folgenden das PG von Bueno de Mesquita näher in den Blick genommen. Da es der Kategorie der nicht-kooperativen Verhandlungsmodelle zuzuordnen ist und in seinen Modellregeln u.a. die Möglichkeit von Konfrontationen zwischen den Entscheidungsakteuren und den Einsatz von Drohmechanismen enthält, kann angenommen werden, dass es besser als die kooperativen Verhandlungsmodelle dazu geeignet sein wird, den Entscheidungsprozess der europäischen Erdgasaußenpolitik gegenüber Russland abzubilden. Dies gilt umso mehr, als es im Test auf den DEUI-Datensatz, der von Bueno de Mesquita im Anschluss an das DEU-Projekt durchgeführt wurde,⁴² insgesamt eine höhere Prognosefähigkeit als die im DEU-Projekt getesteten Modelle aufweist und dies ganz besonders bei politischen Sachfragen, in denen der Konfliktgrad zwischen den EU-Mitgliedstaaten sehr hoch war.

Im Folgenden werden zunächst die Grundzüge des Modells in simplifizierten, nicht-mathematischen Termini skizziert. Daraufhin wird der mathematisch-formale Hintergrund erläutert. Abschließend wird dargelegt, in welcher Form das PG bereits auf den DEUI-Datensatz getestet wurde und gezeigt, dass es besonders akkurate Prognosen bei der Abbildung von EU-Entscheidungsprozessen mit einem hohen Konfliktgrad generiert. Diese Erläuterungen sind für die vorliegende Untersuchung in zweierlei Hinsicht von Relevanz: Zum einen liefern sie eine Grundlage für die Begründung der Modellauswahl in Abschnitt 4.1, die sich daran orientiert, inwiefern von den Spielregeln und der Prognosefähigkeit des Modells *ex ante* die Hypothese abgeleitet werden kann, dass das PG für die Abbildung der europäischen Erdgasaußenpolitik besonders geeignet ist. Zum anderen ist dieses Wissen notwendig, um nach der Übertragung des PG auf die Erdgasaußenpolitik der EU in Kapitel 5 Rückschlüsse von dem Modell auf den wirklichen Entscheidungsprozess ziehen zu können.

42 In seinem aktuellen Entwicklungsstand war das PG nicht Teil des DEU-Projekts. Lediglich eine zuvor entwickelte Version, das „expected utility model“, auf die das PG aufbaut, fand in der Studie von Thomson et al. (2006b) Berücksichtigung (vgl. Arregui et al. 2006). Die erste Version des Modells wurde bereits in den 1980er Jahren von Bueno de Mesquita (1984, 1994, 2002) entwickelt. Darauf aufbauend hat er es im Kontext verschiedener Anwendungen weiterentwickelt. Das PG fügt einen deutlichen höheren Grad an Komplexität zu der Ursprungsversion hinzu und stellt daher eine signifikante Überarbeitung ebendieser dar (vgl. Bueno de Mesquita 2011: 67, 70; Schneider et al. 2011: 11).

Grundzüge des Predictioneer's Game

Das PG, entwickelt von Bruce Bueno de Mesquita (2010, 2011), ist der Kategorie der nicht-kooperativen Verhandlungsmodelle zuzuordnen. Mit der Modellierung des PG ist der Zweck verbunden ein Instrument zu entwickeln, das Zukunftsprognosen für nicht-kooperative Verhandlungssituationen erzeugen kann:

The modeling here is intended to be sufficiently generic that it can be applied to any situation involving the possibility of negotiation in the shadow of the threat (or the realization of the use) of coercion whether in the international arena, the domestic political arena, or in business or social interactions. (Bueno de Mesquita 2011: 66)

Die Regeln des Spiels sind also so allgemein, dass die Simulation auf möglichst viele Konfliktsituationen angewandt werden kann.

Wie bei den Modellierungen des DEU-Projekts handelt es sich beim PG ebenfalls um ein Modell, das die Verhandlungen zu *einer* kontroversen Sachfrage in einem eindimensionalen Raum abbildet (vgl. Bueno de Mesquita o. J.). Mit Bezug auf die Analyse von Entscheidungsprozessen in der EU ist festzuhalten, dass die Simulation die informellen Verhandlungen abbildet, die vor dem formalen Verhandlungsprozess stattfinden, wobei berücksichtigt wird, dass die formalen Vorgaben den informellen Verhandlungsprozess beeinflussen können (vgl. Bueno de Mesquita 1994: 74).

Das Spiel ist iterativ, d.h. es besteht aus mehreren Verhandlungsrunden, deren Anzahl im Vorhinein nicht bestimmt werden kann (vgl. Bueno de Mesquita 2011: 67). Es handelt sich jedoch nicht um ein wiederholtes Spiel wie sie in der Spieltheorie gewöhnlich angewandt werden. Als zentralen Unterschied zu wiederholten Spielen hält Bueno de Mesquita fest, dass sich die Payoffs der Spieler sowie die Werte der Inputvariablen nach jeder Runde in Abhängigkeit von den Gleichgewichten der jeweiligen vorherigen Runde verändern können:

The game is iterated, as distinct from repeated, because payoffs change endogenously (or at least quasi-endogenously, taking both game theoretic and heuristic choices into account) in response to prior stages of play. History, in the shape of dyadic, perfect Bayesian equilibrium outcomes changes the game here whereas repeated games hold payoffs constant, literally repeating interactions over time while allowing for discounting of future values compared to present payoffs. (Bueno de Mesquita 2011: 67)

Weder die Spieler noch der Modellierer wissen im Vorhinein, wann das Spiel endet. Aus diesem Grund muss es simuliert und kann nicht analytisch gelöst werden. Die Ungewissheit über die Anzahl der Verhandlungsrunden bis zur Entscheidungsfindung ist nach Bueno de Mesquita eine wichtige Charakteristik politischer Wirklichkeit und sollte daher in Verhandlungsmodellen, deren Zweck es ist, die wirkliche Welt möglichst adäquat abzubilden, nicht simplifiziert dargestellt werden (vgl. Bueno de Mesquita 2011: 67, 72–73).

Das Spiel kann mit einer unbegrenzten Anzahl an Spielern modelliert werden (vgl. Bueno de Mesquita 2011: 67). Berücksichtigt werden sollen alle Akteure oder Gruppen, die ein Interesse daran haben, das Verhandlungsergebnis zu beeinflussen. Dies muss nicht ausschließlich die Entscheidungsträger umfassen, die aufgrund ihres Amtes an der finalen Entscheidung beteiligt sind, sondern kann auch Akteure betreffen, die die formalen Entscheidungsträger beeinflussen. Die Akteure werden anhand von vier Inputvariablen näher bestimmt: Ihrer Policyposition, die sie bezüglich der politischen Sachfrage einnehmen, ihrem Einfluss auf die Entscheidungsfindung, der Priorität, die sie dem Thema verleihen, und ihrer Kompromissbereitschaft (vgl. Bueno de Mesquita 2010: 50, o. J.). Die Bestimmung der Inputvariablen wird im Folgenden näher erläutert:

- *Policyposition*: Da es sich beim PG um ein eindimensionales räumliches Modell handelt, werden die Positionen in einem Zahlenraum bzw. auf einem Zahlenstrahl eindimensional angeordnet. Letzterer muss für das jeweilige kontroverse Thema zunächst definiert werden. Dazu wird eine Skala mit Werten von 0 bis 100 erstellt, die alle Positionen umfasst, die zu dieser Frage von den in der Simulation berücksichtigten Entscheidungsakteuren eingenommen werden. Bei der Gestaltung des Zahlenstrahls gelten die zuvor erläuterten Anforderungen der räumlichen Modellierung hinsichtlich der Äquivalenz von „politischer Distanz“ und der entsprechenden Entfernung der Policypositionen auf dem Zahlenstrahl sowie der eingipfligen Nutzenfunktionen der Spieler (siehe Abschnitt 2.2). Sofern der Streitgegenstand aus sich heraus eine numerische Unterteilung aufweist, kann diese direkt übernommen werden; ist dies nicht der Fall, müssen die qualitativen Bestimmungen der Positionen in numerische Werte transformiert werden:

The continuum will either have a natural numeric interpretation, such as the percentage of uninsured on health care to be covered under a new policy or the analyst will need to develop numeric values that reflect the relative degree of difference across policy stances that are not inherently quantitative. It is important that the numerical values assigned to different positions (and they can range between any values) reflect the relative distance or proximity of the different solutions to one another. (Bueno de Mesquita o. J.)

Die *Rational Choice*-Modelle im DEU-Projekt, die die Policypositionen der untersuchten Akteure integrieren, unterscheiden sich hinsichtlich der genauen Bestimmung dieser Variable: Manche Modelle fassen darunter die Idealposition der Akteure (siehe z. B. Steuernberg/Selck 2006), andere definieren sie als die tatsächlich geäußerte Verhandlungsposition, mit der die Akteure zu Beginn in die Verhandlungen treten (siehe z. B. Arregui et al. 2006). Letztere kann jedoch von der eigentlichen Idealposition aufgrund von strategischen Überlegungen oder anderen Faktoren abweichen (vgl. Bueno de Mesquita 2004: 130). Zur Verdeutlichung dieser Differenz spricht Thomson (2011) von Positionen und Präferenzen. Der Begriff der Position bezieht sich dabei auf die geäußerte Verhandlungsposition. Die Präferenzen der Akteure sind hingegen versteckt und stimmen nicht notwendigerweise mit den artikulierten Positionen überein (vgl. Thomson 2011: 132). Hält man an dieser Unterscheidung fest, ist es allerdings unmöglich, die Präferenzen eines Akteurs zu ermitteln, da lediglich dem Akteur selbst seine eigenen Präferenzen bekannt sind. Modelle, die in diesem Sinne Präferenzwerte für ihre Analyse benötigen, weisen daher eine unüberwindbare Schwäche hinsichtlich der genauen Bestimmung ihrer Inputvariablen auf. Thomson (2011) weist in seiner DEUII-Studie auf diese Problematik hin, relativiert sie jedoch zugleich, da seine Forschungsergebnisse darauf hindeuten, dass die geäußerten Positionen und die zugrundeliegenden Interessen der Akteure weitgehend kongruent sind und die Positionen die jeweiligen Präferenzen daher ausreichend widerspiegeln, so dass derlei Modellen nicht vorgehalten werden kann, sie seien bloße Gedankenexperimente ohne Wirklichkeitsbezug (vgl. Thomson 2011: 177). Vergleicht man die Prognosefähigkeit der verschiedenen Modelle im DEU-Projekt, sollte diese Unterscheidung jedoch berücksichtigt werden, da die Daten des DEU-Projekts die initiale Verhandlungsposition, nicht

die verdeckten Präferenzen angeben und sofern keine optimalen Werte für Modelle mit präferenzbasierten Policypositionen liefern (vgl. Bueno de Mesquita 2004: 130; Stokman/Thomson 2004b: 13).

Nach Ansicht von Bueno de Mesquita ist es ebenfalls nur sehr bedingt möglich, die Idealpositionen der Akteure zu ermitteln, da davon auszugehen ist, dass sie diese aus strategischen Gründen nicht immer offenbaren (vgl. Bueno de Mesquita 2004: 130). Gleichwohl sieht er durchaus einen wichtigen Beitrag in der Berücksichtigung von Idealpositionen für die spieltheoretische Analyse, sofern ihre Identifizierung möglich ist. Im PG können daher sowohl Idealpositionen als auch die geäußerten Verhandlungspositionen zur Bestimmung der Positionsvariablen in das Modell integriert werden: „The new model can take either ideal points or current bargaining positions. While in practice true ideal points are difficult to know, still this opens the door to a more expansive view of the negotiating process.“ (Bueno de Mesquita 2011: 76).

- *Einfluss*: Diese Variable gibt die Fähigkeit des Akteurs an, die anderen Akteure dazu zu bewegen, sich auf ein Verhandlungsergebnis zu einigen, das der Policyposition dieses Akteurs stärker entspricht. Die Werte für den Einfluss der Akteure sind nicht auf die Skala bis 100 begrenzt; sie müssen über 0 liegen, können aber über den Wert 100 hinausgehen. Eine bewährte Praxis zur Bestimmung der Einflusswerte besteht jedoch darin, dem einflussreichsten Akteur den Wert 100 zuzuschreiben und die Einflusswerte der anderen Akteure in Relation zu diesem festzusetzen (vgl. Bueno de Mesquita 2010: 50, o. J.).
- *Priorität*: Diese Variable beziffert, wie wichtig eine bestimmte politische Sachfrage für einen Akteur ist: Hat sie gegenüber anderen Fragestellungen Priorität? Wie hoch ist der Zeitaufwand, den der Akteur diesem Thema im Vergleich zu anderen Themen widmet? Wie stark engagiert der Akteur sich, um den Entscheidungsprozess zugunsten seiner Position zu beeinflussen? Je höher die Priorität ist, die der Akteur dem Thema verleiht, desto höher ist der Wert dieser Variablen. Er muss jedoch zwischen 0 und 100 liegen. Die Macht eines Akteurs, die sich letztlich im Entscheidungsprozess geltend macht, ist sowohl von seinem Einfluss abhängig als auch von dem Stellenwert, den er dem Thema verleiht und den entsprechenden Einsatz seines Einflusses bedingt. Mathematisch setzt sich die Macht daher aus der Multiplikation von Einfluss und Priorität zusammen (vgl. Bueno de Mesquita 2010: 50, o. J.).

- *Kompromissbereitschaft*: Die Berücksichtigung der Kompromissbereitschaft der Akteure ist eine neu hinzugefügte Variable, die das PG von den vorherigen Entwicklungsstadien des Modells unterscheidet. Die Variable zeigt an, ob es einem Akteur wichtiger ist, an seiner Policyposition festzuhalten und von dieser so wenig wie möglich abweichen zu müssen oder er unbedingt darauf drängt, dass in den Verhandlungen letztlich eine Einigung erzielt wird – auch wenn dies die Bereitschaft erfordern könnte, von seiner Position abzurücken. Die Werte für die Variable liegen zwischen 0 und 100, wobei der Wert 0 anzeigt, dass ein Akteur auf seiner Policyposition beharrt und nicht dazu bereit ist, einem Kompromiss zuzustimmen. Dies mag seiner tatsächlichen Stellung zu den Verhandlungen entsprechen, kann aber auch auf strategischen Überlegungen basieren: „This is, in essence, the extreme view of a true believer or ideologue. Of course, it can also be a bluffed declaration of resolve in an effort to extract larger concessions.“ (Bueno de Mesquita 2011: 75) Unabhängig von dieser Differenzierung kann der Wert im Verlauf des Spiels ansteigen, sofern die Kosten, die aufgebracht werden müssen um die eingangs geäußerte Policyposition zu verteidigen, den Nutzen eines der Policyposition des Akteurs entsprechenden Verhandlungsergebnisses übersteigen. Der Wert 100 zeigt hingegen ein so großes Interesse des Akteurs an einer Einigung der Verhandlungsparteien an, dass er bereit ist, jeglichen Kompromiss zu akzeptieren. Je deutlicher sich diese Variable dem Wert 100 nähert, umso größer ist somit das Interesse des Akteurs an einem Kompromiss und umso geringer das Interesse an dessen inhaltlicher Ausgestaltung. Genauso ist es aber auch möglich, dass die Kompromissbereitschaft eines Akteurs im Verlauf des Spiels sinkt, wenn er aus seinen Erfahrungen mit den Gleichgewichtsergebnissen der vorangegangenen Verhandlungsrunden schließt, dass dies seinen Nutzen steigern würde (vgl. Bueno de Mesquita 2011: 75).

Die notwendigen Informationen für die Inputvariablen erhebt Bueno de Mesquita in der Regel mit Experteninterviews (vgl. Bueno de Mesquita 2010: 50–51).

Nachdem nun erläutert wurde, welche Inputdaten für die Simulation notwendig sind, wird im Folgenden die Spielsequenz skizziert. In dem Spiel treten stets zwei Spieler gegeneinander an. Modelliert man eine Verhandlungssituation mit mehr als zwei Spielern, treffen in der Simulation in jeder Verhandlungsrunde alle direkten Paare gleichzeitig aufeinander (also A vs. B, A vs. C, B vs. C, B vs. A, C vs. B, C

vs. $A, \dots, N - 1$ vs. N , N vs. $N - 1$).⁴³ Das bedeutet, die Verhandlungen laufen für alle Verhandlungspaare parallel ab. Bei N Spielern handelt es sich somit um $N(N - 1)$ Spiele, die in der Simulation gleichzeitig gelöst werden (vgl. Bueno de Mesquita 2011: 72).

Das PG ist ein Spiel mit vollständiger, aber imperfekter Information. Es beginnt mit einem Zug durch die Natur, mit dem die Typen der Spieler bestimmt werden. Dieser Zug entscheidet über zwei Eigenschaften der Spieler: ob sie Falke oder Taube und ob sie friedfertig oder vergeltend sind. Erstere Unterteilung bezieht sich auf Spielsituationen, in denen der zu typisierende Spieler als erster am Zug ist und über die Art des Angebots entscheidet, die er dem Verhandlungspartner vorlegt; letztere Unterteilung gilt für Situationen, in denen der Spieler auf Angebote des Verhandlungspartners reagiert. „Falke“ bezeichnet Spielertypen, die es bevorzugen, den Verhandlungspartner dazu anzuhalten, ihren Forderungen nachzugeben – auch wenn das Kosten für den Verhandlungspartner sowie für den jeweiligen Spieler selbst impliziert – anstatt einen Kompromiss auszuhandeln. Im Gegensatz dazu ziehen „Tauben“ es vor, Kompromisse zu schließen anstatt mittels kostspieliger Pression den Verhandlungspartner zum Nachgeben zu bewegen. Vergeltende Spieler bevorzugen es, sich – möglicherweise zu hohen Kosten – zu verteidigen anstatt zuzulassen, sich von dem Verhandlungspartner auf ein Verhandlungsergebnis drängen zu lassen, das von der eigenen Position abweicht. Ein friedfertiger Spieler zieht es hingegen vor nachzugeben, wenn er von seinem Verhandlungspartner bedrängt wird, um Verteidigungskosten zu vermeiden (vgl. Bueno de Mesquita 2011: 70–71). Die Spieler wissen nicht, welche Kombination von Spielertypen für den jeweils anderen Spieler gilt. Die Wahrscheinlichkeitseinschätzungen liegen zu Beginn des Spiels jeweils bei 0.5 und werden entsprechend der Bayesschen Regel im Verlauf des Spiels angepasst. Bezüglich der Werte der Inputvariablen verfügen die Spieler über vollständige Informationen. Da die Spieler jedoch nicht wissen, welcher Spielertyp der jeweilige Verhandlungspartner ist, können sie die Kosten ihres Gegenübers für dessen mögliche Aktionen nicht berechnen (vgl. Bueno de Mesquita 2011: 70–72).

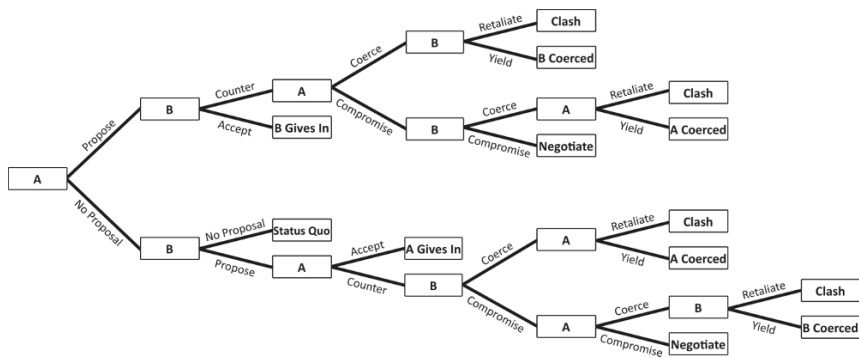
Nach dem Zug der Natur entscheidet der Spieler, der als erster ziehen darf, ob er seinem Verhandlungspartner ein Angebot unterbreitet oder nicht. Ein solches Angebot besteht in der Forderung an den Verhandlungspartner, eine Position

⁴³ Dies impliziert, dass die Spieler nicht wissen, ob sie als erster oder zweiter oder gleichzeitig mit dem anderen Spieler ziehen (vgl. Bueno de Mesquita 2011: 72).

einzunehmen, die für den zuerst ziehenden Spieler vorteilhaft ist. Der Verhandlungspartner hat daraufhin die Möglichkeit nachzugeben und das Angebot anzunehmen oder sich gegen das Angebot zu stellen und dem Anbieter Kosten aufzuerlegen. Die Angebote werden so gewählt, dass sie den erwarteten Nutzen des Spielers am Ende des Spiels maximieren (vgl. Bueno de Mesquita 2010: 245, 2011: 73).⁴⁴

Insgesamt ergibt sich folgende extensive Form einer Verhandlungsrunde für ein Verhandlungspaar:

Abbildung 1: Entscheidungsbaum des Predictioneer's Game



Entscheidungsbaum eines Stufenspiels für ein Verhandlungspaar; Zug der Natur und daraus folgende Informationsmengen werden nicht abgebildet

Quelle: Bueno de Mesquita (2010: 244)

Der Entscheidungsbaum wird in vereinfachter Form dargestellt, da der Zug durch die Natur und die sich daraus ergebenden Informationsmengen nicht abgebildet werden. Eine solche Verhandlungsrunde zwischen zwei Verhandlungspartnern ist zunächst ein in sich abgeschlossenes Spiel, für das ein perfektes Bayesianisches Gleichgewicht ermittelt wird. Wie oben bereits erläutert wurde, besteht die Besonderheit der Simulation aber darin, dass in jeder Runde alle direkten Paare gleichzeitig miteinander verhandeln. Entsprechend werden in der Simulation $(N - 1)^2$ Spiele gleichzeitig gelöst werden. Es werden so viele Runden gespielt, bis eine der beiden Stoppregeln zutrifft. Diese Form der Modellierung ermöglicht es, die Inputvariablen

⁴⁴ Eine detailliertere Erläuterung erfolgt im nächsten Unterkapitel.

der Spieler in Abhängigkeit von den Gleichgewichten der vorangegangenen Runde zu verändern. Die Spieler können also mit einer Policyposition, die von ihrer ursprünglichen Verhandlungsposition abweicht, in die nächste Verhandlungsrunde treten. Sie mögen an Einfluss gewonnen oder verloren haben, dem Verhandlungsgegenstand eine höhere oder niedrigere Priorität zuschreiben oder mehr bzw. weniger Kompromissbereitschaft zeigen als in der vorangegangenen Verhandlungsrunde. Dies hat wiederum maßgeblichen Einfluss auf das Gleichgewicht der neuen Verhandlungsrunde. Auf diese Weise wird in der Simulation berücksichtigt, dass die Spieler nicht in jeder Verhandlungsrunde „bei null“ beginnen, sondern die Entwicklungen in den Verhandlungen stets aufgenommen und fortgeschrieben werden (vgl. Bueno de Mesquita 2011: 72–74).

Das Spiel endet, wenn eine der beiden folgenden Bedingungen erfüllt ist:

- 1) [T]he sum of player payoffs at the end of an iteration is greater than the projected sum of those payoffs in the next iteration, indicating that the average player's welfare is expected to decline in the sense of accumulated payoffs; or
- 2) the sum of player utility, taking into account not only their payoffs from the games in which they are the primary players, but all games including those in which they are third parties, is greater in the current round than the projected sum of utilities in the next iteration, indicating that the average player's welfare is expected to decline in the sense of total utility.

(Bueno de Mesquita 2011: 72)

Die Simulation generiert zwei Spielergebnisse:

The issue forecast takes surround round-by-round predictions into account. The round-by-round predictions are equal to the weighted mean value of the positions of all of the players in that round. The security forecast is the weighted median position of all of the players in the round. The weighted mean or its smoothed version (the issue forecast – my preferred basis for prediction) is the reliable basis for predicting since the player's value for an issue is taken to be two-dimensional (a weighted combination of their policy stance and their eagerness to reach agreement or resist agreement). (Bueno de Mesquita o. J.)

Es handelt sich bei beiden Ergebnissen um sogenannte „point predictions“, die *ein* spezifisches Ergebnis prognostizieren (vgl. Morton 1999: 164). In Anlehnung an Bueno de Mesquita wird in der vorliegenden Arbeit ebenfalls der „issue forecast“,

d.h. der gewichtete Mittelwert der Spielerpositionen in der letzten Verhandlungsrunde als Simulationsergebnis und Zukunftsprognose ausgewählt (vgl. Bueno de Mesquita o. J.).

Nachdem nun die Grundzüge des Modells erklärt worden sind, wird im Folgenden die Berechnung der Payoffs erläutert.

Berechnung der Payoffs

Bueno de Mesquita hat den Code seiner Simulation nicht veröffentlicht.⁴⁵ In seinen Publikationen (2010: 244–247, 2011: 73–75) legt er jedoch die mathematischen Formeln zur Berechnung der Payoffs an den Endnoten eines Stufenspiels zwischen zwei Verhandlungspartnern dar. Diese werden im Folgenden nachgezeichnet.⁴⁶

Wie im vorangegangenen Kapitel erläutert, besteht der erste Zug eines jeden Spielers aus der Entscheidung, dem Verhandlungspartner ein Verhandlungsangebot vorzulegen. Dieses Angebot entspricht nicht zwangsläufig der tatsächlichen Policyposition des Anbieters, sondern liegt meist zwischen den Policypositionen der beiden Verhandlungspartner. Die Angebote werden endogen mit dem Ziel entwickelt, den erwarteten Nutzen des Anbieters zu maximieren:

In practice, this means choosing proposals that make the other players indifferent between imposing costs on the demander and choosing a negotiated compromise instead. A negotiated compromise is always welfare enhancing from the demander's perspective relative to having costs imposed on it by the rival. That is, proposals are chosen to minimize the prospect of being coerced. Of course, the endogenous selection of proposal values must take into account player beliefs about their rival's type. (Bueno de Mesquita 2011: 73)

Die Wahrscheinlichkeit, dass A sich in einer Verhandlungsrunde zwischen A und B durchsetzt, wird folgendermaßen berechnet:

$$P_B^A = \frac{\sum_{\{U(KA)|U(KA)>U(KB)\}} (C_K)(S_K)(U_{KA} - U_{KB})}{\sum_{K=1}^n (C_K)(S_K) | (U_{KA} - U_{KB}) |}$$

45 Scholz, Calbert und Smith (2011) versuchen zwar auf Grundlage einer chronologischen Literaturrecherche das *expected utility model* von Bueno de Mesquita nachzubilden, sie beziehen dabei jedoch nur Publikationen bis zum Jahr 2009 ein.

46 Die Ausführung entspricht den Angaben in Bueno de Mesquita 2011.

mit

$K = 1$ bis n Spieler

C = Einfluss der Spieler

S = Priorität der politischen Sachfrage

U = Nutzen eines Spielers (erstes Subskript) in Bezug auf den Ansatz eines anderen Spielers
(zweites Subskript)

Weitere Definitionen:

$X1_K$ = Policypräferenz von K

$X2_K$ = die Präferenz von K hinsichtlich des Spannungsfelds zwischen dem unbe-
dingten Erzielen einer Einigung sowie dem Beharren auf der eigenen Position

$$U_{SQ}^A = \text{Nutzen von A durch status quo} = \left((1 - (X1_A - X1_{\text{Weighted Mean}})^2) \right) S_A$$

Der Nutzen von A durch Bs Ansatz wird mittels einer *Cobb-Douglas*-Nutzenfunktion
bestimmt:

$$U_B^A = \left[\left(1 - (X1_A - X1_B)^2 \right)^\theta \right] \left[\left(1 - (X2_A - X2_B)^2 \right)^\beta \right]$$

mit $\theta > 0, \beta > 0, \theta + \beta \leq 1$

Anhand der Funktion wird deutlich, dass der Nutzen der Spieler umso größer ist,
je mehr *sowohl* die Kompromissfähigkeit der Spieler *als auch* ihr Ansatz zur Lösung
der politischen Sachfrage übereinstimmen.

Das Modell differenziert zwischen vier Kostenkategorien:

α = die Kosten für den Versuch, den Verhandlungspartner zur Anerkennung des
eigenen Ansatzes zu zwingen und dabei auf Widerstand zu stoßen

τ = die Kosten des eigenen Widerstands gegen die Maßnahmen des Verhandlungs-
partners zur Anerkennung seines Ansatzes

γ = die Kosten, die entstehen wenn der Verhandlungspartner Maßnahmen zur
Durchsetzung seines Ansatzes ergreift und der Spieler keinen Widerstand
leistet

φ = die Kosten für den Versuch, den Verhandlungspartner zur Anerkennung des
eigenen Ansatzes zu zwingen

Die Kosten werden auf Grundlage heuristischer Regeln kalkuliert. Aus der Kategorisierung wird ersichtlich, dass die Kosten in einer Verhandlung zwischen dem Akteur, der seinen eigenen Vorschlag durchsetzen möchte und dem Akteur, der darauf reagiert, differieren. Es wird angenommen, dass die Kosten des Akteurs, der einen Vorschlag unterbreitet, geringer sind als die Kosten des Akteurs, der auf diesen Vorschlag reagieren muss (d.h. $\tau > \alpha$). Des Weiteren besteht die Annahme, dass die antizipierten Kosten der jeweiligen Akteure relativ zur Wahrscheinlichkeit einer Niederlage in den Verhandlungen ansteigen. Die Kosten für Maßnahmen zur Durchsetzung des eigenen Vorschlags ohne Widerstand des Verhandlungspartners werden nicht in Bezug auf die anderen Kostenkategorien definiert, können also größer oder kleiner als diese sein (vgl. Bueno de Mesquita 2011: 74; Bueno de Mesquita/Lalman 1992: 43–44).⁴⁷

Wie zuvor erläutert verändern sich die Werte für die Inputvariablen nach jeder Verhandlungsrunde auf Grundlage heuristischer Regeln unter Berücksichtigung der Gleichgewichte der jeweiligen vorausgegangenen Verhandlungsrunde.

Auf Basis der zuvor aufgezeigten Definitionen bestimmt Bueno de Mesquita die erwarteten Payoffs an den Endknoten eines Stufenspiels. Hier ist erneut zu beachten, dass die Angebote, die die Spieler ihren Verhandlungspartnern machen ($X1_K$), endogen hergeleitet werden und nicht den Policypositionen der Spieler entsprechen müssen.

- Ergebnis 1: As erwarteter Payoff | B akzeptiert As Angebot = $1 - U_{AB}^A$
- Ergebnis 2: As erwarteter Payoff | A ergreift Maßnahmen, um B zur Annahme von As Angebot zu zwingen, B leistet Widerstand = $N_B^A - \alpha_B^A - \varphi_B^A$ (mit N entsprechend der Definition aus Ergebnis 6)
- Ergebnis 3: As erwarteter Payoff | A ergreift Maßnahmen, um B zur Annahme von As Angebot zu zwingen, B willigt ein = $1 - U_{AB}^A - \varphi_B^A$
- Ergebnis 4: As erwarteter Payoff | B ergreift Maßnahmen, um A zur Annahme von Bs Angebot zu zwingen, A leistet Widerstand = $N_B^A - \tau_B^A - \varphi_B^A$

⁴⁷ Bueno de Mesquita und Lalman (1992: 40–41, 43–46) führen die der Kostenkalkulation zugrunde liegenden Annahmen detailliert aus. Sie beziehen sich dabei jedoch auf internationale Konflikte, die potentiell zu kriegerischen Auseinandersetzungen führen können. Eine solche inhaltliche Zuspitzung ist für das PG nicht vorgesehen (vgl. Bueno de Mesquita 2011: 66).

- Ergebnis 5: As erwarteter Payoff | B ergreift Maßnahmen, um A zur Annahme von Bs Angebot zu zwingen, A willigt ein $= 1 - U_{BA}^A - \gamma_B^A$
- Ergebnis 6: As erwarteter Payoff | A und B schließen einen Kompromiss $= (P_B^A) (1 - U_{AB}^A) + (1 - P_B^A) (1 - U_{BA}^A) = N_B^A$
- Ergebnis 7: As erwarteter Payoff | der Status quo zwischen A und B bleibt erhalten $= U_{SQ}^A$
- Ergebnis 8: As erwarteter Payoff | A akzeptiert Bs Angebot $= 1 - U_{BA}^A$
- As erwarteter Nutzen | A bietet an, einen Kompromiss zu schließen:

$$D_B^* N_B^A + (1 - D_B^*) (\operatorname{argmax} [(1 - U_{BA}^A - \gamma_B^A) (N_B^A - \tau_B^A)])$$

- As erwarteter Nutzen | A ergreift Maßnahmen, um B zur Annahme von As Angebot zu zwingen:

$$R_B^* (N_B^A - \alpha_B^A - \varphi_B^A) + (1 - R_B^*) (1 - U_{AB}^A - \varphi_B^A)$$

Mit D^* = Wahrscheinlichkeitseinschätzung, dass der Verhandlungspartner eine Taube ist;
 R^* = Wahrscheinlichkeitseinschätzung, dass der Verhandlungspartner vergeltend ist;
 diese Wahrscheinlichkeitseinschätzungen werden entsprechend der Bayesianischen Regel aktualisiert;

Wahrscheinlichkeitseinschätzungen jenseits des Gleichgewichtspfades werden auf 0.5 gesetzt

Die Angebote werden zwischen den Verhandlungspartnern ausgetauscht. Es sind jedoch nur diejenigen Angebote glaubwürdig, für die eine der beiden folgenden Bedingungen gilt:

- 1) B gibt As Maßnahmen zur Erzwingung von As Vorschlag nach.
- 2) Der absolute Wert eines Angebots abzüglich der gegenwärtigen Position des Angebotsadressaten in Relation zur Bandbreite der möglichen Policydifferenzen ist geringer als der gegenwärtige Wert der Kompromissfähigkeitsvariable des Angebotsadressaten.

Die prognostizierte neue Position eines Spielers am Ende einer Verhandlungsrunde entspricht dem gewichteten Mittelwert der glaubwürdigen Angebote, die er erhält. Das Simulationsergebnis entspricht dem gewichteten Mittelwert⁴⁸ aller glaubwürdigen Angebote der Verhandlungsrunde, in der das Spiel endet. Er wird geglättet als Durchschnitt der gewichteten Mittelwerte – die angrenzenden Verhandlungsrunden direkt vor und nach der Endrunde eingeschlossen.

Anwendung auf den DEU-Datensatz

Wenngleich die Modelle, deren Prognosefähigkeiten im DEU-Projekt getestet wurden, mit dem Ziel entwickelt worden sind, die generellen Merkmale von EU-Entscheidungsprozessen abzubilden, um eine möglichst große Zahl an Fällen analysieren zu können, so sind sie doch für unterschiedliche Kontextbedingungen besser bzw. schlechter geeignet (vgl. Schneider et al. 2011: 8). Als besonders ausschlaggebend hat sich die Differenz zwischen kooperativen und nicht-kooperativen Modellen erwiesen. In dem gemeinschaftlichen Setting der EU sind kooperative Modelle insgesamt besser dazu geeignet, Entscheidungsprozesse abzubilden (vgl. Abschnitt 2.2). Gleichwohl hat das PG bei seiner Anwendung auf den DEU-Datensatz eine hohe Prognosefähigkeit bewiesen. Im Folgenden werden die Ergebnisse dieses Tests näher erläutert.

Das PG wurde erst nach Durchführung der DEUI-Studie entwickelt bzw. veröffentlicht, weshalb nur ein vorheriger Entwicklungsstand, das „expected utility model“, in die DEUI-Studie eingeflossen und seine Prognosefähigkeit mit der der anderen Modelle statistisch verglichen worden ist. Es wies in der DEUI-Studie allerdings eine wesentlich höhere Fehlerrate auf als die am besten prognostizierenden kooperativen Modelle (vgl. Achen 2006a). Dies wurde darauf zurückgeführt, dass die in das Modell integrierten Drohmechanismen nicht in das kooperative Setting der EU passen, da zumindest der regelmäßige Gebrauch von Drohungen in Entscheidungsprozessen in der EU nicht zu erwarten sei, würde dies doch die europäische Gemeinschaft gefährden (vgl. Schneider et al. 2006: 306).

Das PG weist gegenüber dem „expected utility model“ einen wesentlich höheren Grad an Komplexität auf. Um zu überprüfen, ob diese Überarbeitung des Modells dazu beigetragen hat, EU-Entscheidungsprozesse adäquater abzubilden, wurde Bueno de Mesquita der DEUI-Datensatz von Thomson et al. (2006b) bereitgestellt,

48 Der gewichtete Mittelwert entspricht der glaubhaft vorgeschlagenen Policyposition gewichtet durch die Summe aus Einfluss und Priorität (vgl. Bueno de Mesquita 2011: 75).

so dass die Prognosefähigkeit des PG ebenfalls anhand von 162 Fällen getestet werden konnte. Daneben wurde das PG auf einen weniger umfangreichen Datensatz mit neun Fallbeispielen für EU-Entscheidungsprozesse, der Bueno de Mesquita durch Thomson zur Verfügung gestellt wurde, angewandt. Der Vorteil von letzterem Datensatz besteht darin, dass er Werte für die Kompromissbereitschaft der Spieler umfasst, die als Inputvariable im PG erfasst ist. Der DEUI-Datensatz beinhaltet diese Variable hingegen nicht. Da es somit an einer empirischen Grundlage für die Wertbestimmung der Kompromissbereitschaft der Spieler mangelt, legt Bueno de Mesquita als Ausgangspunkt für alle Spieler den Wert 50 fest, der sich genau in der Mitte der Skala befindet. Diese Problematik sollte bei der Beurteilung des Modells berücksichtigt werden (vgl. Bueno de Mesquita 2011: 76).

Um die Prognosefähigkeit des PG unter verschiedenen Kontextbedingungen zu prüfen, gliedert Bueno de Mesquita die 162 Fälle des DEUI-Datensatzes zunächst dahingehend auf, inwiefern sie dem nicht-kooperativen Setting des PG entsprechen. Dazu entwickelte er zwei Kategorien. Zum einen sind in dem Datensatz 125 Fälle enthalten, bei denen bereits eine gemeinsame Position, d.h. ein Status quo, auf den sich die EU in Entscheidungsprozessen zurückbeziehen kann, existiert: „Given the highly cooperative nature of the repeated-play EU decision making environment, these 125 issues are particularly likely to deviate from the non-cooperative expected utility model and the new model [PG, M.G.] because these cases represent repeated games and a likely setting for logrolls.“ (Bueno de Mesquita 2011: 76) Es ist daher zu erwarten, dass die verbleibenden 37 Fälle dem Setting des PG stärker entsprechen. Des Weiteren lassen sich die Fälle hinsichtlich der Frage unterteilen, ob die Diskussionsgegenstände grundlegend neu verhandelt wurden oder lediglich der Status quo modifiziert werden sollte.⁴⁹ Bei den aus dieser Kategorisierung resultierenden 111 Neuverhandlungen geht Bueno de Mesquita ebenfalls davon aus, dass sie in einem relativ kompetitiven Kontext stattfinden (vgl. Bueno de Mesquita 2011: 76–77).

49 Die Kategorien überschneiden sich partiell. Selbstverständlich handelt es sich bei der Mehrheit der Fälle ohne Status quo auch um Neuverhandlungen – nämlich in 29 von 37 Fällen. Gleichwohl gibt es auch Fälle, bei denen sich die EU zu einem früheren Zeitpunkt zwar schon einmal auf eine gemeinsame Position geeinigt hat, die von Thomson et al. (2006b) aber dennoch als Neuverhandlungen eingestuft werden (vgl. Bueno de Mesquita 2011: 77).

Bezüglich der Berechnung des Einflusses der Spieler testet Bueno de Mesquita zwei Varianten. Zum einen testet er das PG mit den Einflusswerten, die er aus der DEUI-Studie übernommen hat: In den Fällen, in denen eine einstimmige Entscheidung der Akteure notwendig ist, wird ihnen der gleiche Einfluss zugeschrieben; in den Fällen, in denen eine qualifizierte Mehrheit ausreicht, wird der Einfluss der Akteure anhand des Shapley-Shubik Index berechnet. Bueno de Mesquita ist jedoch der Ansicht, dass in dem kooperativen Kontext der EU alle Akteure über einen ähnlichen Einfluss verfügen, weshalb er das PG in einem zweiten Test erneut auf den DEUI-Datensatz anwendet, jedoch in allen Fällen eine identische Einflussgröße für die Akteure festlegt (vgl. Bueno de Mesquita 2011: 77).

Um die Prognosefähigkeit des PG in Relation zu den Prognosen der im DEUI-Projekt berücksichtigten Modelle zu testen, bemisst Bueno de Mesquita den *mean absolute percentage error* (MAPE), den *median absolute percentage error* (MdAPE) und die Standardabweichung sowohl für das PG als auch für das expected utility model sowie den Median und einen gewichteten Mittelwert, die sich aus den Inputdaten ergeben. Insbesondere der Mittelwert ist ein geeigneter Vergleichswert, da im DEUI-Projekt kein anderes Modell substantiell bessere Prognosen generiert hat (vgl. Bueno de Mesquita 2011: 77–78).

- *Neun Fälle mit Werten für die Kompromissbereitschaft:* Für die neun von Thomson bereitgestellten Fälle weist das PG die größte Prognosefähigkeit auf. Dies ist unabhängig davon, ob der Einfluss der Spieler immer identisch ist oder – je nach Abstimmungsregel – anhand des Shapley-Shubik Index berechnet wird (vgl. Bueno de Mesquita 2011: 78–79).

Tabelle 1: Fehlerrate formaler Modelle für neun Fälle von Thomson

Modell	MdAPE	MAPE	Standardabweichung
PG, Einfluss ident.	7.7	8.9	8.1
PG, SSI	7.7	8.9	8.1
EUM, Einfluss ident.	18.5	21.5	19.0
EUM, SSI	18.5	21.5	19.0
Median	20.0	29.4	22.7
Mittelwert, Einfluss ident.	12.5	11.8	9.8
Mittelwert, SSI	12.5	11.8	9.8

Quelle: Bueno de Mesquita (2011: 78)

Die Prognosefähigkeit des PG verschlechtert sich jedoch signifikant, wenn keine Werte für die Kompromissbereit der Akteure vorliegen, sondern der Wert für alle Akteure auf 50 festgelegt wird. Unter diesen Bedingungen ist die Prognosefähigkeit des Modells geringer als die des Mittelwerts. Die folgende Tabelle zeigt die Fehlerraten für die divergierenden Werte der Kompromissbereitschaft im Vergleich (vgl. Bueno de Mesquita 2011: 78–79).

Tabelle 2: Fehlerrate des PG mit divergierenden Kompromissbereitschaftswerten

Modell	MdAPE	MAPE	Standardabweichung
PG, Einfluss ident.	7.7	8.9	8.1
PG, Einfluss ident., Kompromiss = 50	16.0	22.2	30.6
PG, SSI	7.7	8.9	8.1
PG, SSI, Kompromiss = 50	10.1	21.9	30.8

Quelle: Bueno de Mesquita (2011: 78)

- *37 Fälle ohne Status quo:* In den Fällen, in denen die EU zuvor noch keine gemeinsame Position erarbeitet hat, weist das PG die höchste Prognosefähigkeit auf, sofern die Einflusswerte der Akteure identisch sind. Mit Einflusswerten, die teilweise anhand des Shapley-Shubik Index berechnet worden sind, liegen

MAPE und Standardabweichung etwas höher als der gewichtete Mittelwert mit identischen Einflusswerten. Obwohl in diesen Fällen keine Werte für die Kompromissbereitschaft der Spieler vorliegen und daher durchgehend auf 50 gesetzt wurden, sind die Fehlerraten des PG hier dennoch wesentlich niedriger als in den neun Fällen von Thomson mit dem Wert 50 für die Kompromissbereitschaft aller Spieler. Diese Ergebnisse lassen darauf schließen, dass das PG gerade in konfliktreichen Entscheidungsprozessen in der EU als Prognoseinstrument gut geeignet ist (vgl. Bueno de Mesquita 2011: 79–80).

Tabelle 3: Fehlerrate formaler Modelle für 37 Fälle ohne Status quo

Modell	MdAPE	MAPE	Standardabweichung
PG, Einfluss ident.	8.2	16.9	24.8
PG, SSI	8.2	19.7	28.6
EUM, Einfluss ident.	10.0	29.4	35.3
EUM, SSI	10.0	28.2	34.7
Median	5.0	19.8	29.8
Mittelwert, Einfluss ident.	8.6	19.4	28.0
Mittelwert, SSI	8.5	19.7	28.6

Quelle: Bueno de Mesquita (2011: 80)

- *111 Fälle mit Neuverhandlungen:* Bei Neuverhandlungen sind die Fehlerraten des PG wesentlich höher als in den Fällen ohne zuvor verhandelten Status quo. Sofern der Einflusswert für alle Spieler identisch festgesetzt wird, ist die Fehlerrate des PG insgesamt aber dennoch am niedrigsten.

Tabelle 4: Fehlerrate formaler Modelle für 111 Fälle mit Neuverhandlungen

Modell	MdAPE	MAPE	Standardabweichung
PG, Einfluss ident.	12.9	23.4	27.0
PG, SSI	18.9	28.0	30.4
EUM, Einfluss ident.	20.0	31.1	32.1
EUM, SSI	24.4	31.3	32.0
Median	20.0	27.9	30.6
Mittelwert, Einfluss ident.	14.5	23.8	27.6
Mittelwert, SSI	16.6	24.7	28.8

Quelle: Bueno de Mesquita (2011: 80)

- *Gesamter DEUI-Datensatz:* Für alle 162 Fälle des DEUI-Datensatzes ist die Fehlerrate des PG geringer als bei der isolierten Betrachtung der Neuverhandlungen. Die Fehlerrate des PG ist – abgesehen vom MdAPE mit gleichen Einflusswerten – etwas höher als die des gewichteten Mittelwerts, aber wesentlich geringer als die des *expected utility model*.

Tabelle 5: Fehlerrate formaler Modelle für vollständigen DEUI-Datensatz

Modell	MdAPE	MAPE	Standardabweichung
PG, Einfluss ident.	12.7	22.8	25.9
PG, SSI	16.2	25.8	28.6
EUM, Einfluss ident.	20.5	30.2	31.1
EUM, SSI	20.5	30.0	31.2
Median	20.0	28.2	30.7
Mittelwert, Einfluss ident.	14.4	22.5	25.5
Mittelwert, SSI	14.0	23.7	27.5

Quelle: Bueno de Mesquita (2011: 81)

Aus dem Vergleich der Fehlerraten des PG bei den unterschiedlichen Zusammensetzungen der Fälle schließt Bueno de Mesquita: „The farther we move from a non-cooperative setting into a more purely cooperative, repeated game environment – such as typifies most European Union decision making – the more likely it is that my iterated, but not repeated game models will fare less well.“ (Bueno de Mesquita 2011: 80)

Insgesamt lassen sich aus den Tests zwei Faktoren ableiten, die den Nutzen des PG für die Analyse von EU-Entscheidungsprozessen maßgeblich beeinflussen: Zum einen ist die Fehlerrate des PG deutlich geringer, wenn Informationen über die Kompromissbereitschaft der Spieler vorliegen. Zum anderen hängt die Prognosefähigkeit des PG vom Setting der Verhandlungen ab. Wenn noch keine gemeinsame Entscheidung für die jeweilige politische Sachfrage getroffen wurde und daher auf ein konfliktives Setting geschlossen werden kann, ist die Fehlerrate des PG geringer als bei der Analyse einer größeren Diversität an Entscheidungsprozessen. Ob es sich um Neuverhandlungen oder die Modifizierung einer bestehenden Einigung handelt, scheint die Prognosefähigkeit des PG hingegen nicht zu beeinflussen. Des Weiteren zeigt der Vergleich mit dem *expected utility model*, dass sich die höhere Komplexität des PG im Sinne einer geringeren Fehlerrate und somit größeren Prognosefähigkeit rentiert.⁵⁰

Bevor in Kapitel 3 die zentralen Konfliktlinien des in dieser Arbeit untersuchten Policy-Subsystems, der europäischen Erdgasaußenpolitik, skizziert und in Abschnitt 4.1 die Auswahl des PG als Analyseinstrument unter Bezugnahme auf die

50 Schneider et al. (2011: 11–12) betonen ebenfalls die geringere Fehlerrate des PG im Vergleich zu den in der DEUI-Studie getesteten Modellen, verweisen aber zugleich auf eine von Schneider et al. (2010a) entwickelte Variante der Nash'schen Verhandlungslösung, die eine noch geringere Fehlerrate (MAE 19.49) für den vollständigen DEUI-Datensatz aufweist (vgl. Schneider et al. 2010a: 97). Dazu ist allerdings anzumerken, dass für das Modell von Schneider et al. nur der MAEP-Wert angegeben wird. Es wäre daher interessant, auch die Standardabweichung sowie den MdAEP-Wert für die Modelle zu vergleichen, um weitere Informationen zu erhalten, wie signifikant die Unterschiede hinsichtlich der Prognosefähigkeit sind (vgl. Bueno de Mesquita 2011: 81). Zudem sollten die Modelle unter verschiedenen Kontextbedingungen getestet werden. Da es sich bei der Variante der Nash-Lösung um ein kooperatives Modell handelt, ist zu erwarten, dass das Modell von Schneider et al. in einem Datensatz mit ausschließlich kooperativen Fällen eine noch geringere Fehlerrate aufweist, wohingegen das PG in nicht-kooperativen Settings voraussichtlich treffendere Prognosen generiert als die Nash-Lösung.

zuvor präsentierten Ergebnisse zu dessen Prognosefähigkeit im EU-Kontext begründet werden, wird im Folgenden die methodische Vorgehensweise dieser Untersuchung ausführlich erläutert.

2.4 Methodische Vorgehensweise dieser Arbeit

Nach der theoretischen Einführung in Forschungsstand, Anwendungsbereiche und Spannungsfelder des Simulationsansatzes in der Politikwissenschaft wird nun die konkrete methodische Vorgehensweise der vorliegenden Arbeit erläutert. Ziel der folgenden Untersuchung ist es, eine Simulation zu erstellen, die die europäische Erdgasaußenpolitik gegenüber Russland möglichst genau abbildet und die Erstellung von Zukunftsszenarien für dieses Policy-Subsystem ermöglicht. Den darauf aufbauenden Anwendungsfall stellt eine Simulation des Wandels in der zukünftigen EU-Erdgasaußenpolitik gegenüber Russland infolge der Ukraine-Krise dar. Die Arbeit gliedert sich folglich in zwei aufeinander aufbauende Stufen:

- 1) Das Modell wird in Form einer Nachbildung der vergangenen EU-Erdgasaußenpolitik gegenüber Russland entwickelt und einem *performance test* unterzogen (*Gas Game I*).
- 2) Das Modell wird auf Grundlage von empirischen Daten aus dem ersten Jahr der Ukraine-Krise modifiziert und eine Prognose über die zukünftige Entwicklung der EU-Erdgasaußenpolitik gegenüber Russland generiert (*Gas Game II*). Aus dem Vergleich der Simulationsergebnisse von vergangener und zukünftiger Erdgasaußenpolitik (*Gas Game I* und *II*) lassen sich Hypothesen hinsichtlich der Frage ableiten, inwiefern die Ukraine-Krise eine Zäsur in der EU-Erdgasaußenpolitik gegenüber Russland darstellt und ein Politikwandel zu erwarten ist.

Im Folgenden wird die methodische Vorgehensweise zur Umsetzung der beiden Schritte ausgeführt.

2.4.1 *Gas Game I: Post-diction* der EU-Erdgasaußenpolitik vor der Ukraine-Krise

Der erste Schritt für die Entwicklung einer anwendungsbezogenen Simulation besteht darin, auf Grundlage der Erfahrungen mit der Wirklichkeit entweder Annahmen über den zu simulierenden Gegenstand aufzustellen und diese in Modellregeln zu transformieren oder auf bereits entwickelte Modelle, von denen man annimmt, dass

sie das entsprechende Wirklichkeitssegment adäquat abbilden, zurückzugreifen. In dieser Arbeit wird mit dem PG ein Modell ausgewählt, das schon sehr erfolgreich für die Simulation von Entscheidungsprozessen in der EU angewandt wurde (vgl. Abschnitt 2.3). Eine ausführliche Begründung für die Auswahl des Modells erfolgt in Abschnitt 4.1.

Durch die Wahl eines bereits entwickelten Modells sind die Regeln der Simulation schon vorgegeben. Um das PG an das jeweilige Politikfeld bzw. Policy-Subsystem – in dieser Arbeit die europäische Erdgasaußenpolitik gegenüber Russland – anzupassen, müssen die relevanten Entscheidungsakteure ausgewählt werden. Zur näheren Bestimmung der Akteure beinhaltet das PG – wie im vorangegangenen Abschnitt beschrieben – vier Inputvariablen: die Policyposition, den Einfluss, die Priorität des Themas und die Kompromissfähigkeit der Akteure. Da das Modell für Zukunftsszenarien eingesetzt werden soll, ist eine enge Anlehnung an die wirkliche Welt erforderlich. Die Inputdaten für die Variablen werden daher in der Wirklichkeit erhoben und diese realen Daten daraufhin für die Modellvariablen operationalisiert.⁵¹

Um zu überprüfen, ob die Modellregeln sowie die Inputdaten geeignet sind, um die Erdgasaußenpolitik der EU gegenüber Russland möglichst adäquat abzubilden, ist der Blick in die Vergangenheit notwendig. Anwendungsbezogene spieltheoretische Modelle werden in der Regel mittels sogenannter „post-dictions“ (Bueno de Mesquita 2004: 127) entwickelt und getestet:

[B]y ‘forecasting’ we mean predictions over periods of time during which the variables are regarded to show changes in value, such that prediction requires that we predict the changes. One form of such forecasting is ‘postdiction’, when information taken from what was once a ‘past’ stratum is used to predict what then was in the present. [...]his type may be used to test dynamic models. (Wayman 2014: 6)

Der von Wayman angesprochene Test der Modelle erfolgt durch einen Vergleich von Simulationsergebnis und realer Politik. *Post-dictions* sind also

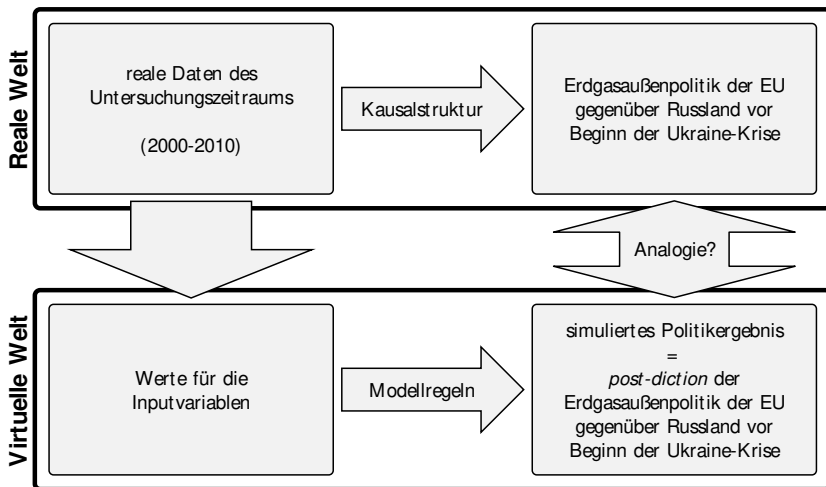
51 Eine Darstellung der Erhebung und Operationalisierung der Inputdaten erfolgt in den jeweiligen Unterkapiteln zu den vier Variablen Policyposition (Abschnitt 4.3), Einfluss (Abschnitt 4.4), Priorität (Abschnitt 4.5) und Kompromissbereitschaft (Abschnitt 4.6).

conditional statements about a phenomenon for which the researcher actually has data, i.e. the outcome (or dependent) variable has been observed, but when making the prediction the researcher pretends as if the values of the dependent variable were unknown. Thus, the prediction can instantly be compared with what has been observed. (Bechtel/Leuffen 2010: 311)

Zur Entwicklung des *Gas Game I*, die in Abbildung 2 illustriert ist, werden dementsprechend Daten aus der Vergangenheit – konkret zwischen 2000 und 2010 – erhoben und in numerische Werte für die Inputvariablen in der Simulation transformiert. Begründet ist der Untersuchungszeitraum darin, dass in dieser Periode zum einen die Debatten zu den Pipelineprojekten Nord Stream, South Stream und Nabucco zwischen den relevanten Entscheidungsakteuren stattfanden, die als empirische Grundlage für die Erhebung der Policypositionen dienen und zum anderen für die Inputvariablen, die mittels quantitatitiven Daten erhoben werden, Durchschnittswerte ermittelt werden können, die die materielle Versorgungslage der Entscheidungsakteure über die gesamte Zeitspanne des Entscheidungsprozesses hinweg widerspiegeln. Mit dem generierten Simulationsergebnis wird der Anspruch verbunden, die europäische Erdgasaußenpolitik gegenüber Russland vor Beginn der Ukraine-Krise im November 2013 *ex post* „vorherzusagen“. Es wird hinsichtlich dieses Zwecks im Rahmen eines *performance tests* mit der wirklichen Welt verglichen. Besteht eine Analogie zwischen *post-diction* und realer Erdgasaußenpolitik im November 2013, generieren reale und virtuelle Welt also identische Policy-Outputs, ist eine erste Bedingung erfüllt, um davon auszugehen, dass die Modellregeln der Simulation die Kausalstruktur der Wirklichkeit adäquat abbilden (vgl. Abschnitt 2.1).

Mit der Entwicklung und Überprüfung von Modellen mittels *post-dictions* sind jedoch zwei Probleme verbunden. Zum einen müssen die Werte für die Inputvariablen die Situation zu *Beginn der Verhandlungen* widerspiegeln, schließlich wird der anschließende Verhandlungsprozess darauf aufbauend simuliert. Das Wissen um das tatsächliche Politikergebnis birgt allerdings die Gefahr, die retrospektive Bestimmung der ursprünglichen Inputwerte zu beeinflussen und zu verfälschen. Bueno de Mesquita erläutert dieses Problem der „post-dictive bias“ (zit. n. Veen 2011a: 25) am Beispiel der Festlegung der Policypositionen der jeweiligen Akteure zu Beginn des Entscheidungsprozesses mittels Experteninterviews:

Abbildung 2: *Post-diction* der EU-Erdgasaußenpolitik gegenüber Russland vor Beginn der Ukraine-Krise im November 2013



Quelle: Eigene Darstellung

If the input data on positions reflect the experts' view of initial positions tainted by ex post knowledge of where each stakeholder ended up, then models that correctly predict a dynamic process will prove less accurate in post-dictive studies than in predictive investigations, because the input data will misrepresent the stage at which the model enters the analytic process. This is thus far an unexamined potential empirical disadvantage arising from post-diction as compared with prediction. (Bueno de Mesquita 2004: 129, siehe auch 1997; Schneider et al. 2006; Veen 2011a)

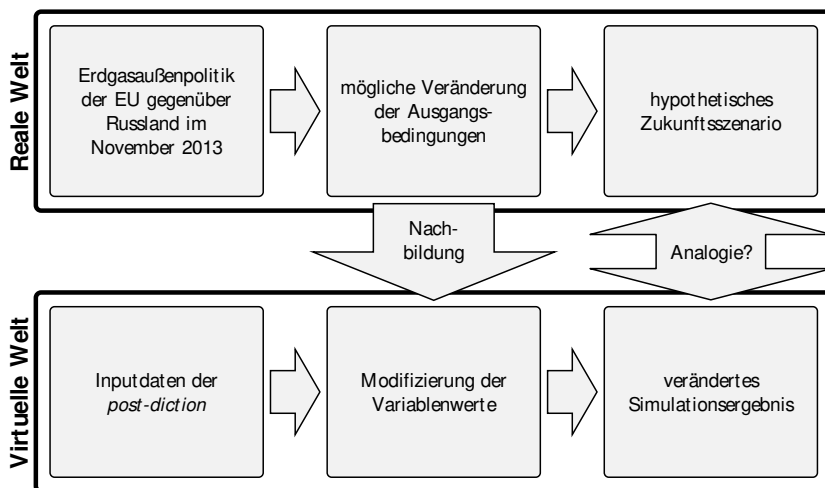
Zwar lassen sich auch Studien finden, die diese Annahme widerlegen (siehe z. B. Rojer 1999); gleichwohl verdeutlichen die skizzierten Überlegungen, dass bei der Bestimmung der Werte für die Inputvariablen zumindest ein hoher Reflexionsgrad des Wissenschaftlers im Forschungsprozess erforderlich ist.

Die zweite Schwierigkeit ergibt sich daraus, dass es sich bei einer *post-diction* lediglich um *einen* Fall handelt. Es könnte auch ein Zufall sein, dass Simulationsergebnis und Wirklichkeit in diesem Fall übereinstimmen. Zwar ist die Konkordanz

von *post-diction* und gegenwärtiger Politik eine notwendige Bedingung, man kann sich aber dennoch nicht darauf verlassen, dass die Simulation auch in anderen Fällen Ergebnisse generiert, die in der Wirklichkeit in ähnlicher Form auftreten.

Aus diesen beiden Gründen ist es notwendig, die Validität des Modells zu überprüfen, bevor es dazu eingesetzt wird, Zukunftsszenarien zu erstellen. Die in dieser Arbeit durchgeführte Validitätsprüfung setzt am zweiten Problem an: So wurde festgehalten, dass eine mit der Wirklichkeit vereinbare *post-diction* ein notwendiger, aber kein hinreichender Hinweis auf die Analogie von Simulation und Wirklichkeit ist. Die Stärke von Simulationen besteht aber darin, dass mit ihnen verschiedene Szenarien erstellt werden können. Daher kann die Plausibilität des Modells zusätzlich getestet werden, indem man weitere Szenarien durchspielt und prüft, ob in jenen Szenarien ebenfalls Simulationsergebnisse produziert werden, die reale Politikergebnisse widerspiegeln. Dieser Vorgehensweise ist zweifellos entgegenzusetzen, dass es nur *eine* wirkliche europäische Erdgasaußenpolitik gibt, mit der das Simulationsergebnis verglichen werden kann. Es ist aber möglich, sich auf Grundlage des bestehenden Wissens über das Policy-Subsystem – die Forschung zur Erdgasaußenpolitik der EU ist sehr umfassend – weitere Zukunftsszenarien für die reale Welt zu überlegen und zu testen, ob die Simulation dazu geeignet ist, diese Szenarien nachzubilden. Die entsprechende Verfahrensweise wird in Abbildung 3 illustriert: Ausgehend vom Status der EU-Erdgasaußenpolitik in der wirklichen Welt vor Beginn der Ukraine-Krise wird die Frage gestellt, wie sich die EU-Erdgasaußenpolitik sehr wahrscheinlich verändern würde, wenn sich bestimmte Ausgangsbedingungen ändern. Es werden somit hypothetische – aber auf Grundlage wissenschaftlicher Expertise zu erwartende – Zukunftsszenarien erstellt. Daraufhin erfolgt der Rückbezug auf das Modell, indem die veränderten Bedingungen der „realen Welt“ in der Simulation durch eine Modifizierung der Daten nachgebildet werden. Schließlich werden die hypothetischen Zukunftsszenarien der „realen Welt“ mit den jeweiligen Simulationsergebnissen verglichen und geprüft, ob die Simulation auch diese Szenarien adäquat abbildet.

Sofern die Simulation all diese Bedingungen erfüllt, ist davon auszugehen, dass sie sehr gut dazu geeignet ist, die europäische Erdgasaußenpolitik gegenüber Russland abzubilden und somit Rückschlüsse aus der Analyse der virtuellen Welt auf die wirkliche Welt ermöglicht. Die Entwicklung des Modells sowie sein *performance test* wären damit erfolgreich abgeschlossen.

Abbildung 3: Validitätsprüfung des *Gas Game I*

Quelle: Eigene Darstellung

2.4.2 *Gas Game II: Prediction* der EU-Erdgasaußenpolitik nach der Ukraine-Krise

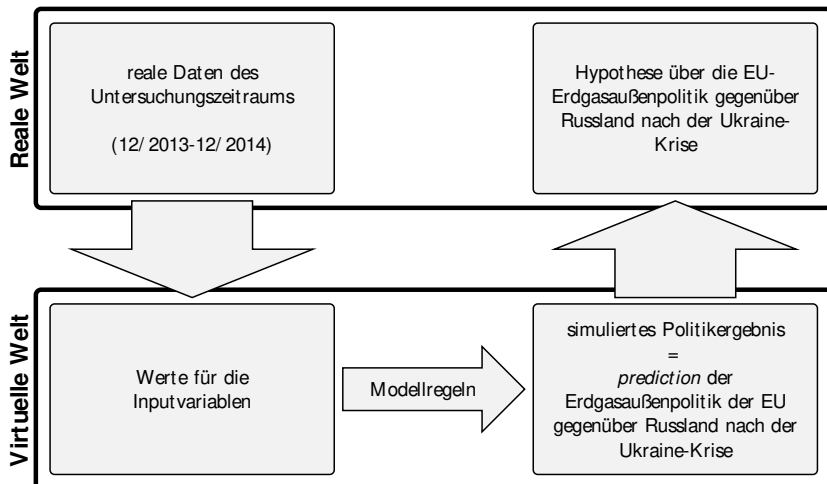
Die auf der Modellentwicklung und den *performance test* aufbauende Stufe besteht darin, eine Prognose über die zukünftige Entwicklung der europäischen Erdgasaußenpolitik gegenüber Russland zu erstellen. Vor dem Hintergrund des Abbruchs des South Stream-Projekts durch den russischen Präsidenten Putin und seiner Kennzeichnung der Entscheidung als Reaktion auf die europäische Energiepolitik wird die Hypothese aufgestellt, dass die Ukraine-Krise eine Zäsur in der europäischen Erdgasaußenpolitik gegenüber Russland darstellt. Mittels einer Modifizierung der Inputvariablen des *Gas Game I*, die die Veränderungen in der realen Welt im ersten Jahr der Ukraine-Krise abbildet, soll ein Simulationsergebnis generiert werden, das einer auf dem Modell basierenden Erwartung über die Entwicklung der europäischen Erdgasaußenpolitik gegenüber Russland infolge der Ukraine-Krise entspricht. Um die Übertragbarkeit des *Gas Game* auf den neuen Untersuchungszeitraum zu gewährleisten, müssen zunächst zwei Bedingungen überprüft werden:

- 1) Bewegt sich der während der Ukraine-Krise geführte Diskurs um die Ausgestaltung der zukünftigen europäischen Erdgasaußenpolitik in demselben politischen Raum wie der analysierte Diskurs des vergangenen Jahrzehnts?
- 2) Sind die Kausalstrukturen der Wirklichkeit unverändert geblieben, so dass eine weitere Analogie zwischen virtueller und wirklicher Welt angenommen werden kann?

Zur Überprüfung der ersten Bedingung werden die zentralen Konfliktlinien zwischen den Entscheidungsakteuren während der Ukraine-Krise identifiziert und hinsichtlich ihrer Identität mit den Konfliktlinien des vergangenen Jahrzehnts abgeglichen. Zur Überprüfung der zweiten Bedingung werden die Positionen der Entscheidungsakteure zur Energieunion, dem Kommissionprojekt zur Entwicklung einer kohärenten europäischen Energiepolitik, analysiert. Aufbauend auf diese Tests wird die modifizierte Simulation wie in Abbildung 4 veranschaulicht erstellt: Die Erhebung der empirischen Daten und deren Transformation in numerische Werte für die Inputvariablen erfolgt analog zu der Vorgehensweise beim *Gas Game I* für den Untersuchungszeitraum von Dezember 2013 bis Dezember 2014, dem ersten Jahr der Ukraine-Krise. Die Simulation generiert mit diesen Werten ein Ergebnis, das eine *prediction* der EU-Erdgasaußenpolitik gegenüber Russland darstellt. Aus dem Simulationsergebnis können Rückschlüsse auf die reale Welt gezogen und Erwartungen über die zukünftige EU-Erdgasaußenpolitik formuliert werden. Des Weiteren werden auf Grundlage eines Vergleichs der Simulationsergebnisse von *Gas Game I* und *Gas Game II*, der in Abbildung 5 illustriert ist, Hypothesen über einen zukünftigen Policy-Wandel in der europäischen Erdgasaußenpolitik entwickelt.

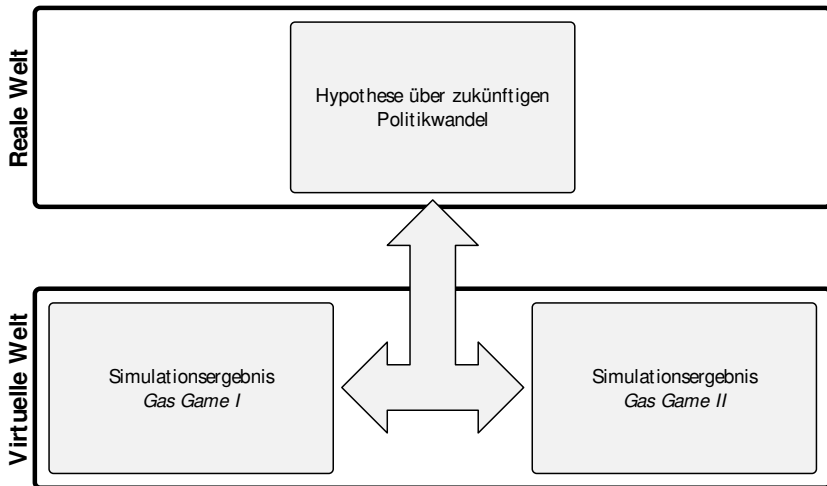
Nachdem die methodische Vorgehensweise skizziert worden ist, werden im Folgenden die zuvor erläuterte Entwicklung des Modells sowie sein *performance test* durchgeführt.

Abbildung 4: *Prediction* der EU-Erdgasaußenpolitik gegenüber Russland nach der Ukraine-Krise



Quelle: Eigene Darstellung

Abbildung 5: Hypothesengenerierung zum Wandel in der EU-Erdgasaußenpolitik gegenüber Russland nach der Ukraine-Krise



Quelle: Eigene Darstellung

Gas Games

Der Wandel der europäischen Erdgasaußenpolitik
infolge der Ukraine-Krise

Grabau, M.

2018, XIII, 615 S. 58 Abb., Softcover

ISBN: 978-3-658-20154-8