

Diskussionspapier des  
Instituts für Organisationsökonomik

2/2018

**Pläne als konditionale Strategien**

Ein Konzept für mögliche Kooperation im Gefangenendilemma

Alexander Dilger

Discussion Paper of the  
Institute for Organisational Economics

**Diskussionspapier des  
Instituts für Organisationsökonomik  
2/2018**

Februar 2018

ISSN 2191-2475

**Pläne als konditionale Strategien  
Ein Konzept für mögliche Kooperation im Gefangenendilemma**

*Alexander Dilger*

**Zusammenfassung**

In der traditionellen Spieltheorie sind Strategien äquivalent mit Handlungen und gemischte Strategien die einzige Erweiterung dazu. Als eine neue Erweiterung werden Strategien als Pläne interpretiert. Obgleich die meisten Pläne nicht sehr interessant sind (wie eine bestimmte Handlung zu tun), sind manche bedingten Pläne es doch. So können sie beispielsweise Kooperation im Gefangenendilemma ermöglichen.

JEL-Codes: C72, C70, D83

# **Plans as Conditional Strategies**

## **A Concept Enabling Cooperation in the Prisoners' Dilemma**

### **Abstract**

In traditional game theory, strategies are equivalent with actions and mixed strategies are the only extension. As a new extension, strategies are interpreted as plans. Although most plans are not very interesting (like doing a certain action), some conditional plans are. For example, they enable cooperation in the prisoners' dilemma.

Im Internet unter:

[http://www.wiwi.uni-muenster.de/io/forschen/downloads/DP-IO\\_02\\_2018](http://www.wiwi.uni-muenster.de/io/forschen/downloads/DP-IO_02_2018)

Westfälische Wilhelms-Universität Münster  
Institut für Organisationsökonomik  
Scharnhorststraße 100  
D-48151 Münster

Tel: +49-251/83-24303 (Sekretariat)  
E-Mail: [io@uni-muenster.de](mailto:io@uni-muenster.de)  
Internet: [www.wiwi.uni-muenster.de/io](http://www.wiwi.uni-muenster.de/io)

# Pläne als konditionale Strategien

## Ein Konzept für mögliche Kooperation im Gefangenendilemma\*

### 1. Einleitung

In diesem Beitrag werden Pläne als ein neues (zumindest teilweise, siehe III.) spieltheoretisches Konzept vorgeschlagen, welches das bekannte Konzept der Strategien erweitert. Diese Erweiterung ist intuitiv plausibel und umfasst die bisherige Spieltheorie als Spezialfall. In gewisser Weise ist der Strategiebegriff in der bisherigen Spieltheorie unpassend, weil es sich nicht um Strategien im ursprünglichen Wortsinne handelt, sondern eigentlich um die Handlungen, aus denen ein Spieler wählen kann. Eine reine Strategie ist dann eine ausgewählte Handlung, eine gemischte Strategie (siehe von Neumann/Morgenstern 1944, S. 146, oder Nash 1950) eine Lotterie mit bestimmten Wahrscheinlichkeiten für verschiedene Handlungen, während in der extensiven Spielform eine Strategie eine Funktion ist, die Handlungen jedem beliebigen bisherigen Handlungsverlauf zuordnet (vgl. Osborne/Rubinstein 1994, S. 92). Außerhalb der Spieltheorie ist eine Strategie hingegen mehr als eine Handlung, nämlich ein großer und langfristiger Plan zum Erreichen grundlegender Zielen (vgl. Chandler 1962). Eine Taktik ist hingegen ein kleinerer und kurzfristigerer Plan über geeignete Mittel zum Erreichen von bestimmten Zwecken bzw. Zwischenzielen. Schließlich gibt es neben der strategischen und taktischen Ebene noch die operative, auf der konkrete Handlungen ausgewählt und ausgeführt werden. Im militärischen Bereich würden sich vor allem die Generale um die Strategie kümmern, die Offiziere um die Taktik und die Soldaten um das Operative.

Als neues spieltheoretisches Konzept sollen nun Pläne vorgestellt werden. Dabei kommt es auf den Namen nicht an. Sie könnten auch Strategien heißen, wenn man für diese dann einen anderen Namen findet, oder z. B. konditionale bzw. bedingte Strategien. Denn es geht darum, dass nicht einfach nur Handlungen ausgewählt oder randomisiert werden können, sondern diese auf alles Mögliche bedingt werden können, nicht nur vergangene Handlungen. Natürlich sind auch einfache Pläne ohne jede Bedingung möglich, die dann einfach den Handlungen bzw. reinen Strategien entsprechen. Ein Plan kann auch vorsehen, verschiedene Handlungen mit unterschiedlichen Wahrscheinlichkeiten durchzuführen, was dann einer gemischten Stra-

---

\* Dieser Beitrag wurde zuerst auf Englisch bei GAMES 2008, dem dritten Weltkongress der Game Theory Society in Evanston, Illinois, am 13. Juli 2008 präsentiert. Es folgte eine zweite Präsentation auf Deutsch beim Ausschuss Wirtschaftswissenschaften und Ethik des Vereins für Socialpolitik in Wittenberg am 1. März 2013. Ein geplanter Sammelband dazu ist bis heute nicht erschienen. Die englische Version dieses Beitrags wurde bereits als Diskussionspapier des Instituts für Organisationsökonomik 5/2014 veröffentlicht, im Internet unter [https://www.wiwi.uni-muenster.de/io/de/forschen/downloads/DP-IO\\_05\\_2014](https://www.wiwi.uni-muenster.de/io/de/forschen/downloads/DP-IO_05_2014).

ategie entspricht. Interessant wird es jedoch dann, wenn Pläne auf etwas anderes konditioniert werden, insbesondere andere Handlungen und Pläne.

Im nachfolgenden 2. Abschnitt werden einige einfache Beispiele für Pläne vorgestellt. 3. behandelt den eigentlich interessanten Fall, dass die Pläne auf fremde Handlungen oder sogar Pläne konditioniert werden. Im 4. Abschnitt wird das Konzept der Pläne exemplarisch auf das Gefangenendilemma angewandt, in dem dadurch Kooperation möglich wird. Unter 5. werden mögliche Probleme des Konzepts diskutiert. Dieser Beitrag schließt mit einem kurzen Fazit im 6. Abschnitt.

## 2. Beispiele für Pläne

Ein ganz einfaches Beispiel für einen Plan ist die Absicht eines Spielers, an Montagen Handlung A durchzuführen und an allen anderen Wochentagen Handlung B. Dieser Plan bezieht sich, wie alle Pläne, auf (intendierte) Handlungen, aber er ist bedingt auf etwas anderes als reine Wahrscheinlichkeiten sowie gemischte Strategien oder auf vergangene Handlungen. Die Bedingung ist hier der Wochentag, an dem zu handeln ist. Offensichtlich kann die Existenz solch eines Plans die Gleichgewichte eines Spiels verändern.

Ein etwas anspruchsvolleres Beispiel lässt sich am Koordinationsspiel „Kampf der Geschlechter“ (Battle of the Sexes, siehe Luce/Raiffa 1957, S. 90-94) entwickeln. Bei diesem Spiel wollen ein (männlicher) Spieler und eine Spielerin gemeinsam etwas unternehmen, wobei der Spieler ein Treffen im Fußballstadion (Handlung A) bevorzugen würde und die Spielerin einen Kinobesuch (Handlung B). Beide sind nicht fest an einem Ort verabredet und enttäuscht, wenn sie sich verfehlen. In traditioneller Form ist dieses Spiel in Abbildung 1 dargestellt.

		Spielerin	
		Strategie A <sub>2</sub> = Handlung A	Strategie B <sub>2</sub> = Handlung B
Spieler	Strategie A <sub>1</sub> = Handlung A	2, 1	-1, -1
	Strategie B <sub>1</sub> = Handlung B	-1, -1	1, 2

**Abbildung 1: Kampf der Geschlechter, traditionelle Form**

In dieser Form hat das Spiel drei Nash-Gleichgewichte. Es können beide A wählen oder beide B. Außerdem gibt es ein Gleichgewicht in gemischten Strategien, bei dem der (männliche) Spieler A mit der Wahrscheinlichkeit  $3/5$  und B mit der Gegenwahrscheinlichkeit  $2/5$  wählt, während die Spielerin es genau umgekehrt macht, also mit der Wahrscheinlichkeit  $2/5$  A wählt und mit der Gegenwahrscheinlichkeit von  $3/5$  B. Dieses dritte Gleichgewicht ist das am wenigsten nutzenstiftende für beide, während der Spieler das erste Gleichgewicht mit beiderseitigem A befürwortet und die Spielerin das zweite mit beiderseitigem B.

Es wird nun der erste Plan  $M_1$  eingeführt, nämlich dass der Spieler an Montagen A wählt und an allen anderen Wochentagen B, wie in Abbildung 2 dargestellt. Es verbleiben dann die alten Gleichgewichte, während es kein neues Gleichgewicht gibt, welches konstant an allen Wochentagen Bestand hätte. Allerdings stellt an Montagen die Kombination  $M_1$  und  $A_2$  ein zusätzliches Gleichgewicht dar, während an den übrigen Wochentagen  $M_1$  und  $B_1$  ein neues Gleichgewicht bilden. Entsprechend gibt es auch neue Gleichgewichte in gemischten Strategien in Abhängigkeit vom Wochentag.

Allerdings hat die Spielerin keine beste Antwort auf Plan  $M_1$ , solange sie selbst keinen Plan in Abhängigkeit vom Wochentag einführt, sondern bei traditionellen Handlungsstrategien verbleibt. Ansonsten ist die beste Antwort auf  $M_1$  ein entsprechender Plan  $M_2$ , an Montagen A zu wählen und von Dienstag bis Sonntag B. In Abbildung 3 ist diese Erweiterung dargestellt mit  $M_1$  und  $M_2$  als neuem Gleichgewicht an jedem Wochentag, allerdings mit anderen Handlungen und Nutzenwerten an Montagen als an anderen Wochentagen.

		Spielerin	
		Strategie $A_2$ = Plan $A_2$ : Handlung A in jedem Fall	Strategie $B_2$ = Plan $B_2$ : Handlung B in jedem Fall
Spieler	Strategie $A_1$ = Plan $A_1$ : Handlung A in jedem Fall	2, 1	-1, -1
	Strategie $B_1$ = Plan $B_1$ : Handlung B in jedem Fall	-1, -1	1, 2
	Plan $M_1$ : Handlung A an Montagen, ansonsten Handlung B	2, 1 an Montagen, -1, -1 an allen anderen Tagen	-1, -1 an Montagen, 1, 2 an allen anderen Tagen

**Abbildung 2: Kampf der Geschlechter mit einem zusätzlichen Plan**

		Spielerin		
		Strategie A <sub>2</sub> = Plan A <sub>2</sub> : Handlung A in jedem Fall	Strategie B <sub>2</sub> = Plan B <sub>2</sub> : Handlung B in jedem Fall	Plan M <sub>2</sub> : Handlung A an Montagen, ansonsten Handlung B
Spieler	Strategie A <sub>1</sub> = Plan A <sub>1</sub> : Handlung A in jedem Fall	2, 1	-1, -1	2, 1 an Montagen, -1, -1 an allen ande- ren Tagen
	Strategie B <sub>1</sub> = Plan B <sub>1</sub> : Handlung B in jedem Fall	-1, -1	1, 2	-1, -1 an Montagen, 1, 2 an allen anderen Tagen
	Plan M <sub>1</sub> : Handlung A an Mon- tagen, ansonsten Handlung B	2, 1 an Mon- tagen, -1, -1 an allen anderen Tagen	-1, -1 an Mon- tagen, 1, 2 an allen ande- ren Tagen	2, 1 an Montagen, 1, 2 an allen anderen Tagen

**Abbildung 3: Kampf der Geschlechter mit zusätzlichem Plan für beide Geschlechter**

Natürlich gibt es unendlich viele von solchen Plänen und Gleichgewichten. Man braucht die Handlungen nur auf Dienstag statt Montage zu konditionieren oder auf 8 Uhr morgens, Regen vor dem Fenster, einen Ort in Australien etc. Allerdings erscheinen solche Pläne und Gleichgewichte nicht besonders wichtig und rechtfertigen für sich genommen noch nicht die Erweiterung der traditionellen Spieltheorie um Pläne.

### 3. Pläne konditioniert auf Handlungen und Pläne

Es gibt mindestens eine Klasse von Bedingungen, die zu neuen Ergebnissen führt und Pläne zu einer interessanten und relevanten Erweiterung von Strategien im Sinne von Handlungen führt. Diese Bedingungen sind solche, die sich auf Handlungen und Pläne beziehen und nicht nur auf externe Ereignisse und Umstände wie den Wochentag. Statt M<sub>1</sub> und M<sub>2</sub> aus Abbildung 3 könnten Spielerin und Spieler jeweils einen Plan E haben, genau die Handlung des anderen zu wählen, wie in Abbildung 4 dargestellt.

		Spielerin		
		Strategie A <sub>2</sub> = Plan A <sub>2</sub> : Handlung A in jedem Fall	Strategie B <sub>2</sub> = Plan B <sub>2</sub> : Handlung B in jedem Fall	Plan E <sub>2</sub> : Hand- lung A, wenn der Spieler A aus- führt, ansonsten Handlung B
Spieler	Strategie A <sub>1</sub> = Plan A <sub>1</sub> : Handlung A in jedem Fall	2, 1	-1, -1	2, 1
	Strategie B <sub>1</sub> = Plan B <sub>1</sub> : Handlung B in jedem Fall	-1, -1	1, 2	1, 2
	Plan E <sub>1</sub> : Handlung A, wenn die Spielerin A ausführt, ansonsten Handlung B	2, 1	1, 2	?

**Abbildung 4: Kampf der Geschlechter mit Plänen bedingt auf die Handlung des anderen**

Allerdings zeigt die Kombination der Pläne E<sub>1</sub> und E<sub>2</sub> ein Problem von Plänen, welches in dieser Form nicht für Handlungen bzw. Handlungsstrategien existiert. Es ist möglich, dass Pläne, die auf andere Pläne konditioniert sind, nicht wohldefiniert sind. Das erkannte bereits Nigel Howard, der aufbauend auf eine Idee von von Neumann/Morgenstern (1944) Meta-spiele mit Metastrategien konstruierte, die eine wohldefinierte Teilmenge der hier vorgeschlagenen Pläne sind (siehe Howard 1966a, 1966b, 1971 und 1976 sowie Rapoport 1967 und für eine Kritik Harris 1969). Die Pläne E<sub>1</sub> und E<sub>2</sub> legen zusammen nicht fest, ob beide die Handlung A oder B ausführen werden. Ihre Pläne oder zumindest der Plan eines Spielers müssen etwas differenzierter sein, um zu einem klaren Ergebnis zu kommen. Der Spieler könnte z. B. statt E<sub>1</sub> den Plan S<sub>1</sub> verfolgen, wonach er A tut, wenn die Spielerin sich unbedingt für A entscheidet, und sonst B. Die Kombination S<sub>1</sub> und E<sub>2</sub> führt dann dazu, dass beide B tun, wie in Abbildung 5 dargestellt.

		Spielerin		
		Strategie A <sub>2</sub> = Plan A <sub>2</sub> : Handlung A in jedem Fall	Strategie B <sub>2</sub> = Plan B <sub>2</sub> : Handlung B in jedem Fall	Plan S <sub>2</sub> : Hand- lung A, wenn der Spieler A unbe- dingt ausführt, ansonsten Hand- lung B
Spieler	Strategie A <sub>1</sub> = Plan A <sub>1</sub> : Handlung A in jedem Fall	2, 1	-1, -1	2, 1
	Strategie B <sub>1</sub> = Plan B <sub>1</sub> : Handlung B in jedem Fall	-1, -1	1, 2	1, 2
	Plan S <sub>1</sub> : Handlung A, wenn die Spielerin A unbedingt ausführt, ansonsten Handlung B	2, 1	1, 2	1, 2

**Abbildung 5: Kampf der Geschlechter mit jeweils einem differenzierteren Plan**

Wenn die Spielerin einen entsprechenden Plan S<sub>2</sub> hat, A zu tun, wenn der Spieler A unbedingt tut, und sonst B, dann führt die Kombination der Pläne S<sub>1</sub> und S<sub>2</sub> zu demselben Ergebnis wie die Pläne S<sub>1</sub> und E<sub>2</sub>, nämlich dass beide B tun werden. Beides sind Gleichgewichte. Natürlich gibt es noch unendlich viel mehr mögliche Pläne, von denen einige Paare keine definierte Lösung haben so wie E<sub>1</sub> und E<sub>2</sub>. Viele weitere führen nicht zu einem Gleichgewicht. Schließlich ist festzustellen, dass das ursprüngliche Koordinationsproblem vom Kampf der Geschlechter auch durch Pläne nicht wirklich gelöst wird. Denn welches Gleichgewicht von Plänen auszuwählen ist, bleibt unbestimmt, während es viel mehr Gleichgewichte gibt als in der traditionellen Spieltheorie, aus denen ausgewählt werden könnte.

#### 4. Pläne im Gefangenendilemma

Wie schon zuvor bei den Plänen bedingt auf Wochentage oder vergleichbare externe Umstände, so wäre auch allein für Spiele wie den Kampf der Geschlechter die Einführung von Plänen über Handlungen und gemischte Strategien hinaus nicht hinreichend interessant. Doch es gibt andere Spiele, bei denen Pläne einen echten und großen Unterschied machen. In diesem Beitrag soll als Beispiel das Gefangenendilemma untersucht werden, in dem Kooperation nicht

nur möglich, sondern sogar zum bevorzugten Gleichgewicht für wirklich rationale Spieler wird.

Ein traditionelles Gefangenendilemma (siehe z. B. Luce/Raiffa 1957, S. 94-97, mit anderen Auszahlungen bzw. Nutzenwerten) kann z. B. um die naiven Pläne N ergänzt werden, immer dann zu defektieren (D), wenn der andere Spieler unbedingt defektiert (was die dominante Strategie in der normalen spieltheoretischen Analyse ohne Pläne ist), und andernfalls zu kooperieren (C). Dies wird in Abbildung 6 dargestellt.

		Spielerin		
		Strategie C <sub>2</sub> = Plan C <sub>2</sub> : C (Kooperieren) in jedem Fall	Strategie D <sub>2</sub> = Plan D <sub>2</sub> : D (Defektieren) in jedem Fall	Plan N <sub>2</sub> : D, wenn der Spieler D unbedingt ausführt, ansonsten C
Spieler	Strategie C <sub>1</sub> = Plan C <sub>1</sub> : C (Kooperieren) in jedem Fall	2, 2	0, 3	2, 2
	Strategie D <sub>1</sub> = Plan D <sub>1</sub> : D (Defektieren) in jedem Fall	3, 0	1, 1	1, 1
	Plan N <sub>1</sub> : D, wenn die Spielerin D unbedingt ausführt, ansonsten C	2, 2	1, 1	2, 2

**Abbildung 6: Gefangenendilemma mit naiven Plänen für bedingtes Kooperieren**

Wenn beide Spieler nur zwischen diesen drei Plänen C, D und N wählen könnten, würde es zwei Gleichgewichte geben statt nur eines ohne das Konzept von Plänen und damit ohne Plan N. Die Kombination D<sub>1</sub> und D<sub>2</sub> bleibt natürlich ein Gleichgewicht. Aber die Kombination N<sub>1</sub> und N<sub>2</sub> ist nun ein weiteres Gleichgewicht, welches für beide Spieler besser ist. Deshalb sollten rationale Spieler eigentlich dieses Gleichgewicht wählen (solange keine weiteren Strategien zugelassen werden).

Das wirkt jedoch seltsam. Das Gefangenendilemma ist das Standardbeispiel für die individuelle Rationalität eines kollektiv irrationalen bzw. ineffizienten Verhaltens. Es hat bereits

viele ungültige Versuche gegeben, die das Gegenteil zeigen sollten. Allerdings macht das Konzept der Pläne bedingte Kooperation möglich. Sobald ein Spieler unter der Bedingung kooperieren kann, dass auch der andere Spieler kooperiert, wird diese Form der bedingten Kooperation zu einer rationalen Option, die auch noch besser ist als wechselseitiges Defektieren. Das ist durchaus bekannt und nichts Neues (siehe z. B. Shubik 1970). Ein bindender Vertrag oder irgendein anderer Durchsetzungsmechanismus mittels einer dritten Partei, etwa ein ungeschriebenes Gesetz der Mafia, jeden Verräter zu bestrafen, ist eine Standardlösung bzw. ein Ausweg aus dem Gefangenendilemma, indem die Struktur der Spiels und die Auszahlungen verändert werden.

Der Unterschied beim Konzept der Pläne ist das Fehlen einer solchen dritten Partei, deren Existenz und Anreize oft zweifelhaft erscheinen, wodurch das Dilemma bestehen bleibt. Pläne können effektiv sein, indem sie die Form eines impliziten Vertrages annehmen, der sich selbst durchsetzt. Die Strafe fürs Defektieren ist das sofortige Defektieren durch den anderen Spieler, während bei wiederholten Spielen ein Bestrafen durch Defektieren in späteren Runden möglich ist, was jedoch bei nicht (potentiell) unendlich häufiger Wiederholung zum Problem der letzten Runden führt, nach der ein Bestrafen nicht mehr möglich ist, was durch Rückwärtsinduktion auf alle vorherigen Runden ausstrahlt (siehe z. B. Selten 1978 oder Milgrom/Roberts 1982, S. 283, und für Gegenargumente Pettit/Sudgen 1989 und Sobel 1993). Das sofortige Bestrafen wird möglich durch die Bedingtheit von Plänen. Nach Plan N wird (unbedingte) Defektion mit Defektion beantwortet und Kooperation mit Kooperation.

Was ist, wenn der Spieler zu kooperieren plant oder das vielleicht auch nur geschickt vor-täuscht, um dann im letzten Moment doch noch zu defektieren? Das scheint die bestmögliche Option für ihn zu sein, die Spielerin durch Täuschung zum Kooperieren zu bringen, um dann selbst durch Defektion besonders viel zu gewinnen. Doch wenn der Spieler von Anfang an einen solchen Wechsel plant, dann ist und war sein Plan von Anfang an zu defektieren, worauf die Spielerin gemäß ihrem Plan N ebenfalls mit Defektion reagiert. Doch selbst wenn der Spieler ursprünglich zu kooperieren plante und dann später seinen Plan änderte, ist das Resultat dasselbe, da Plan N nicht bedingt ist auf den Zeitpunkt vom Plan des Spielers oder gar auf dessen ersten Plan, sondern auf seinen effektiven Plan bzw. die tatsächlich daraus folgende Handlung.

Die Pläne  $N_1$  und  $N_2$  in Abbildung 6 stellen nicht wirklich ein Gleichgewicht dar und wirklich rationale Spiele benötigen ausgefeiltere Pläne. Denn  $N_1$  ist nicht die beste Antwort auf  $N_2$  (und natürlich ebenso umgekehrt). Gegeben  $N_2$  könnte der Spieler durch einen Wechsel von

$N_1$  zu einem anderen bedingten Plan gewinnen, z. B. zu  $N_1^+$ : „Kooperiere, wenn die Spielerin unbedingt kooperiert, ansonsten defektiere.“ In dem Fall defektiert der Spieler nicht unbedingt, so dass die Spielerin entsprechend Plan  $N_2$  kooperiert, doch zugleich kann der Spieler auf diese Kooperation mit Defektion antworten. Das ist natürlich nicht im Interesse der Spielerin, die dementsprechend einen anderen Plan verfolgen sollte. Natürlich wären auch  $N_1^+$  und ein entsprechender Plan  $N_2^+$  kein Gleichgewicht.

Ein Gleichgewicht besteht aus wechselseitig besten Antworten, so dass kein Spieler einen Anreiz hat, seine Strategie bzw. seinen Plan zu ändern. Dies bedeutet, dass im Gefangenendilemma beide Spieler ihre eigene Kooperation davon abhängig machen müssen, dass der jeweils andere Spieler kooperiert, und sie nicht von etwas anderem abhängig machen dürfen, so dass der andere Spieler trotz Defektion die Vorteile der eigenen Kooperation genießen kann. Am einfachsten erscheint der Plan  $O_1$ : „Kooperiere, wenn die Spielerin kooperiert, ansonsten defektiere.“ Doch die Plankombination  $O_1$  mit dem korrespondierenden Plan  $O_2$  führt zu demselben Problem wie in Abbildung 4, dass die Handlungen indeterminiert sind, weil beide Spieler gemeinsam sowohl kooperieren als auch defektieren könnten.

Ein ausgeklügelterer bedingt kooperativer Plan ist wie in Abbildung 7 dargestellt  $P_1$ : „Kooperiere, wenn die Spielerin nur unter der Bedingung kooperiert, dass ich selbst kooperiere, ansonsten defektiere.“ Die Spielerin kann den symmetrischen Plan  $P_2$  haben: „Kooperiere, wenn der Spieler nur unter der Bedingung kooperiert, dass ich selbst kooperiere, ansonsten defektriere.“ Die Strategiekombination  $P_1$  und  $P_2$  stellt ein kooperatives Gleichgewicht dar, weil die Bedingungen erfüllt sind, so dass beide kooperieren und keiner einen Anreiz hat, seinen Plan zu ändern. Allerdings ist Symmetrie natürlich nicht erforderlich für ein Gleichgewicht.  $P_1$  und  $O_2$  ergeben z. B. ein kooperatives Gleichgewicht mit asymmetrischen Plänen.

		Spielerin		
		Strategie C <sub>2</sub> = Plan C <sub>2</sub> : C (Kooperieren) in jedem Fall	Strategie D <sub>2</sub> = Plan D <sub>2</sub> : D (Defektieren) in jedem Fall	Plan P <sub>2</sub> : C, wenn der Spieler C nur unter dieser Bedin- gung ausführt, ansonsten D
Spieler	Strategie C <sub>1</sub> = Plan C <sub>1</sub> : C (Koope- rieren) in jedem Fall	2, 2	0, 3	0, 3
	Strategie D <sub>1</sub> = Plan D <sub>1</sub> : D (Defek- tieren) in jedem Fall	3, 0	1, 1	1, 1
	Plan P <sub>1</sub> : C, wenn die Spiele- rin C nur unter die- ser Bedingung aus- führt, ansonsten D	3, 0	1, 1	2, 2

**Abbildung 7: Gefangenendilemma mit bedingt kooperativem Gleichgewicht**

## 5. Mögliche Probleme und Diskussion

Die skizzierte Kooperationsmöglichkeit im Gefangenendilemma durch bedingte Pläne wirft weitere Fragen auf. Woher kennen z. B. die Spieler die Pläne des jeweils anderen und die sich daraus ergebenden Handlungen? Allerdings ist dies ein Problem in der Spieltheorie allgemein, also auch in der traditionellen Variante, nicht nur beim neuen Konzept der Pläne. Um z. B. beim Kampf der Geschlechter eines der beiden traditionellen Gleichgewichte in reinen Strategien zu erreichen, muss man die Strategie bzw. Handlung des anderen Spielers kennen oder zumindest erraten. Das ändert nichts daran, dass A die beste Antwort ist, wenn der andere Spieler bzw. die Spielerin auch A wählt. Entsprechend ist  $N_1$  die beste Antwort auf  $N_2$  und umgekehrt.

Allerdings gibt es eine zusätzliche Schwierigkeit bei den hier skizzierten Plänen. So hängt nicht nur der zu wählende Plan vom Plan des anderen Spielers ab, sondern auch die auszuführende Handlung innerhalb dieses Plans. In der traditionellen Spieltheorie gibt es keinen wesentlichen Unterschied zwischen Strategien und Handlungen, jedenfalls nicht für reine Strategien, während gemischte Strategien zusätzlich nur vom Zufall abhängen, nicht von der

Strategie des anderen Spielers. Nichtsdestotrotz erfordert das Handeln im Rahmen eines Plans kein anderes Wissen als die Auswahl eines Plans, der ein Gleichgewicht darstellt. Schlimmstenfalls ist es eben nicht möglich, einem bestimmten Plan zu folgen. Wenn man gar keine Idee (keinen Plan) hat, welchem Plan der andere Spieler folgen könnte, dann könnte ein Plan wie  $N$  schlicht nicht durchführbar sein. Faktisch ist es dann so, als wenn der Plan nicht existieren würde und der Raum möglicher Pläne entsprechend beschränkt wäre, im Extremfall auf reine Handlungen und deren zufällige Mischung wie in der traditionellen Spieltheorie. Das bedeutet jedoch nicht, dass der Raum möglicher Pläne stets so eingeschränkt ist. Die traditionelle Spieltheorie ist nur ein Spezialfall der viel größeren Möglichkeiten, die Pläne bieten.

Es ist weder irrational noch unmöglich, Pläne zu haben, die von den Handlungen oder Plänen anderer abhängen. In der Theorie können Modelle konstruiert werden, für die dies einfach angenommen wird. Es sind auch Hybridmodelle möglich und vielleicht sogar von besonderem Interesse, in denen einige Spieler die Pläne von anderen korrekt (oder mit einer bestimmten Wahrscheinlichkeit) vorhersagen können, während andere Spieler diese Fähigkeit nicht haben, sie aber in anderen erkennen können, während noch andere Spieler das Konzept von Plänen rigoros ablehnen und nur die traditionelle Spieltheorie anwenden. Mit letzteren ist eine Kooperation im Gefangenendilemma nicht möglich, zwischen Spielern aus den ersten beiden Gruppen hingegen schon.

Ob reale Menschen Pläne von anderen Menschen erkennen oder zumindest hinreichend gute Abschätzungen vornehmen können, ist eine empirische Frage. Die Antwort ist höchstwahrscheinlich positiv, zumindest für manche Menschen unter bestimmten Umständen. Anonymes Spielen z. B. über Computer könnte dafür jedoch unzureichend sein. Vermutlich müssen die Spieler zuerst miteinander reden oder sich auch in die Augen schauen, bevor sie bedingte Pläne entwerfen und konkrete Handlungen wählen können, die auch die Handlungen und Pläne ihrer Mitspieler berücksichtigen.

In vielen Versionen des Gefangenendilemmas ist vorherige Kommunikation verboten, was jedoch eine unnötige Restriktion darstellt. Die Logik des Dilemmas wird durch unverbindliches Reden („cheap talk“) nicht verändert, solange die Entscheidungen simultan und vor allem unbeobachtet vom jeweils anderen Spieler erfolgen. Dagegen könnte solche unverbindliche Kommunikation im Vorfeld sehr wertvoll sein, um Pläne zu koordinieren und Erwartungen über den Plan des Mitspielers zu entwickeln oder einfach um den Typ des Mitspielers zu bestimmen, z. B. welche Vorstellungen von Rationalität und welche Kooperationsbereit-

schaft er oder sie hat. Entsprechend könnte es für dieselben Spieler rational sein, ohne vorherige Kommunikation zu defektieren und mit bzw. nach ihr zu kooperieren.

Allerdings setzt die Fähigkeit, nach solchen bedingten Plänen zu handeln, etwas mehr voraus als reine Rationalität, nämlich die Fähigkeit, die Pläne anderer Spieler zu erkennen oder zumindest mit hinreichender Genauigkeit vorherzusagen. Dann ist es nicht irrational, sondern gerade ein Zeichen von besonderer Rationalität, sich auf Kooperation zu einigen und eine entsprechende Einigung auch einzuhalten. Zugleich muss man vorsichtig bleiben, damit man nicht einer nicht vertrauenswürdigen Person vertraut. Solches Vertrauen einschließlich Vertrauenswürdigkeit haben in der traditionellen Spieltheorie keinen Platz oder wären dort sogar irrational. Wenn der andere Spieler nur an die traditionelle Spieltheorie glaubt, ist in der Tat Defektieren die beste und einzig sinnvolle Option, im Übrigen in Übereinstimmung mit Plan N.

Beim endlich oft wiederholten Gefangenendilemma ist die Interaktion zwischen einem Anhänger der traditionellen Spieltheorie, Spieler T, und einer Spielerin P, die das Konzept von Plänen akzeptiert, interessanter. Zwei Spieler vom Typ T könnten nur die ganze Zeit defektieren, wie allgemein bekannt ist, während zwei Spielerinnen vom Typ P die ganze Zeit kooperieren könnten einschließlich der letzten Runde, da sie auch ohne Wiederholung zu kooperieren vermögen (siehe unter 4.). Ohnehin ist Kooperation in vielen Runden leichter als in nur einer, weil man den Mitspieler oder die Mitspielerin kennenlernt und auch wechselseitig Reputation aufgebaut werden kann. Ob Spielerin P auch in den meisten Runden mit Spieler T kooperieren kann, hängt von noch einer weiteren Eigenschaft neben ihrer Rationalität und Fähigkeit zum Erkennen von Plänen bzw. in diesem Fall des Typs T ab, aus dem alle Handlungen deterministisch folgen. Spielerin P bräuchte zusätzlich die Fähigkeit, sich an einen Plan selbst binden zu können. Wenn Spielerin P sich erkennbar bindet, in jeder Runde einschließlich der letzten zu kooperieren, solange Spieler T zuvor kooperiert hat, kann auch Spieler T bis auf die letzte Runde kooperieren. Spieler T wird zwar Spielerin P für irrational halten, doch eigentlich ist Spielerin P rationaler und steigert die Auszahlungen für sich (und Spieler T) ungemein.

## **6. Fazit**

Pläne, die auf andere Pläne und Handlungen bedingt sind, erlauben Kooperation im Gefangenendilemma. Natürlich bleibt wechselseitiges Defektieren ein Gleichgewicht. Dieses ist jedoch schlechter für beide Spieler als ein Gleichgewicht mit wechselseitig bedingter Koope-

ration. Defektieren ist nicht länger eine dominante Strategie wie in der traditionellen Spieltheorie mit reinen Handlungsstrategien oder auch ein dominanter Plan, da es ohnehin keine nichttrivialen dominanten Pläne mehr gibt. Triviale Dominanz würde bedeuten, dass ein Spieler indifferent ist zwischen allen Alternativen, die die anderen Spieler herbeiführen können. Wenn das nicht der Fall ist, dann gibt es für jeden beliebigen Plan Q einen Plan Anti-Q von dem oder den Mitspielern, der dazu führt, dass der Spieler mit Plan Q das schlechtestmögliche Ergebnis unter Q erhält (im Gefangenendilemma wäre das wechselseitiges Defektieren) und bei jedem anderen Plan das bestmögliche Ergebnis (im Gefangenendilemma würde der Wechsel von Q mit Kooperation vom Mitspieler belohnt, selbst wenn der Wechsel seinerseits mit Defektieren verbunden wäre). Die Existenz von Anti-Q zerstört jede nichttriviale Dominanz im Sinne einer einzigen besten Antwort auf alle Pläne des anderen Spielers.

Umgekehrt gibt es nicht nur unendlich viele verschiedene Pläne ohne klare Ordnung, sondern es gibt auch stets beliebig viele Gleichgewichte. Deren Existenz (immer) und Eindeutigkeit (nie) sind keine relevanten Fragen im Kontext von Plänen. Interessanter ist die Suche nach Effizienz oder z. B. Robustheit von Plan-Gleichgewichten. In jedem Fall ist bedingte Kooperation ein vorteilhafter und vernünftiger Plan im Gefangenendilemma oder bei vergleichbaren Kollektivgutproblemen mit mehr Spielern.

## Literatur

- Chandler, Alfred D. (1962): „Strategy and Structure: Chapters in the History of the American Industrial Enterprise“, MIT Press, Cambridge, MA.
- Harris, Richard J. (1969): „Note on Howard’s Theory of Meta-Games“, Psychological Reports 24 (3), S. 849-850.
- Howard, Nigel (1966a): „The Theory of Meta-Games“, General Systems 11, S. 167-186.
- Howard, Nigel (1966b): „The Mathematics of Meta-Games“, General Systems 11, S. 187-200.
- Howard, Nigel (1971): „Paradoxes of Rationality: Theories of Metagames and Political Behavior“, MIT Press, Cambridge, MA.
- Howard, Nigel (1976): „Prisoner’s Dilemma: The Solution by General Metagames“, Behavioral Science 21 (6), S. 524-531.
- Luce, Robert Duncan/Raiffa, Howard (1957): „Games and Decisions: An Introduction and Critical Survey“, John Wiley & Sons, New York.
- Milgrom, Paul/Robert, John (1982): „Predation, Reputation, and Entry Deterrence“, Journal of Economic Theory 27 (2), S. 280-312.

- Nash, John F. (1950): „Equilibrium Points in n-Person Games“, Proceedings of the National Academy of Sciences of the United States of America 36 (1), S. 48-49.
- Osborne, Martin J./Rubinstein, Ariel (1994): „A Course in Game Theory“, MIT Press, Cambridge, MA.
- Pettit, Philip/Sugden, Robert (1989): „The Backward Induction Paradox“, Journal of Philosophy 86 (4), S. 169-182.
- Rapoport, Anatol (1967): „Escape from Paradox“, Scientific American 217 (2), S. 50-56.
- Selten, Reinhard (1987): „The Chain-Store Paradox“, Theory and Decision 9 (2), S. 127-159.
- Shubik, Martin (1970): „Game Theory, Behavior, and the Paradox of the Prisoner’s Dilemma: Three Solutions“, Journal of Conflict Resolution 14 (2), S. 181-193.
- Sobel, Jordan Howard (1993): „Backward-Induction Arguments: A Paradox Regained“, Philosophy of Science 60 (1), S. 114-133.
- von Neumann, John/Morgenstern, Oskar (1944): „Theory of Games and Economic Behavior“, Princeton University Press, Princeton.

## Diskussionspapiere des Instituts für Organisationsökonomik

Seit Institutsgründung im Oktober 2010 erscheint monatlich ein Diskussionspapier. Im Folgenden werden die letzten zwölf aufgeführt. Eine vollständige Liste mit Downloadmöglichkeit findet sich unter <http://www.wiwi.uni-muenster.de/io/de/forschen/diskussionspapiere.html>

- DP-IO 2/2018** Pläne als konditionale Strategien  
Ein Konzept für mögliche Kooperation im Gefangenendilemma  
*Alexander Dilger*  
Februar 2018
- DP-IO 1/2018** Öffentliche Betriebe  
*Alexander Dilger*  
Januar 2018
- DP-IO 12/2017** Verzerrungen bei Personalbeurteilungen durch Führungskräfte  
*Julia Müller*  
Dezember 2017
- DP-IO 11/2017** Kommerzieller Organhandel aus ökonomischer Sicht  
*Alexander Dilger*  
November 2017
- DP-IO 10/2017** 7. Jahresbericht des Instituts für Organisationsökonomik  
*Linn-Brit Bakkenbüll/Alexander Dilger*  
Oktober 2017
- DP-IO 9/2017** Sind gewählte Teamleiter besser als ihr Team?  
Analyse am Beispiel von BWL-Professoren und ihren Zitationen  
*Alexander Dilger*  
September 2017
- DP-IO 8/2017** Zur Stärkung wissenschaftlicher Kritik  
*Alexander Dilger*  
August 2017
- DP-IO 7/2017** Theoretische Erklärungsansätze für die Entsprechenserklärungen zu Abfindungen für Vorstandsmitglieder  
*Ute Schottmüller-Einwag*  
Juli 2017
- DP-IO 6/2017** Doping in Teams  
A Simple Decision Theoretic Model  
*Alexander Dilger*  
Juni 2017
- DP-IO 5/2017** Wirtschaftsethische Überlegungen zur Finanz- und Eurokrise  
*Alexander Dilger*  
Mai 2017
- DP-IO 4/2017** Vor- und Nachteile verschiedener Arten von Drittmitteln  
*Alexander Dilger*  
April 2017
- DP-IO 3/2017** The Euro from a Business Perspective  
*Alexander Dilger*  
März 2017



Herausgeber:  
Prof. Dr. Alexander Dilger  
Westfälische Wilhelms-Universität Münster  
Institut für Organisationsökonomik  
Scharnhorststr. 100  
D-48151 Münster

Tel: +49-251/83-24303

Fax: +49-251/83-28429

[www.wiwi.uni-muenster.de/io](http://www.wiwi.uni-muenster.de/io)

