

Fairness und Reziprozität im Diktatorspiel

Axel Franzen und Sonja Pointner

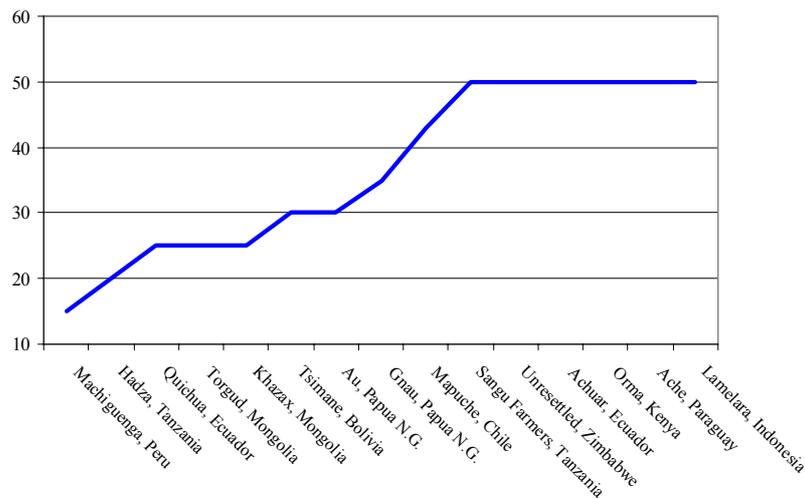
1. Einführung

In der Spieltheorie sind in den letzten Jahren zunehmend verhaltenstheoretische Ansätze in den Vordergrund gestellt worden, die sich mit Abweichungen vom ökonomischen rationalen Entscheidungsmodell beschäftigen. Das ökonomische Standardmodell kann das Verhalten in vielen Situationen nicht erklären. Die Ergebnisse in Diktator- und Ultimatumspielen deuten darauf hin, dass in den Verhaltensweisen von Menschen neben egoistischen Motiven Altruismus, Fairness und Reziprozität eine große Rolle spielen (vgl. Diekmann 2004; Fehr, Gächter 2000; Ockenfels 1999). In vielen Fällen orientieren sich die Akteure nicht am Nash-Gleichgewicht, sondern die Spieler weichen zum Teil erheblich von rationalen Strategien ab (Forsythe u.a. 1994). Das Ausmaß der Abweichung variiert dabei in verschiedenen experimentellen Studien (z.B. Henrich et al. 2004). Angenommen wird, dass das Vorhandensein von sozialen Normen wesentlich ist für das Zustandekommen von Kooperation. Warum diese Normen Kooperation bedingen können ist allerdings nicht das Thema des Aufsatzes.

In diesem Beitrag steht das Aufzeigen der Existenz der Normen Fairness, Reziprozität und Altruismus mit Hilfe des Diktatorspiels im Vordergrund. Die Entstehung dieser Normen ist keine Selbstverständlichkeit, sondern eine enorme zivilisatorische Leistung. Viele indigene Völker haben diese Normen nicht im selben Maß ausgeprägt wie dies in westlichen Industrie-

gesellschaften der Fall ist. Die Ergebnisse zum Ultimatumspiel veranschaulichen dies. Im Ultimatumspiel gibt es ein „Verhandlungsszenario“: Ein Spieler (Proposer) kann eine fixe Summe aufteilen, jedoch hat der Mitspieler (Responder) die Möglichkeit, das Angebot abzulehnen. Nimmt er das Angebot an, bekommen beide Spieler die vereinbarte Summe, lehnt er ab, so erhält niemand etwas. Die Prognose gemäß dem Standardmodell ist, dass der Proposer nicht mehr als die kleinste mögliche Einheit anbieten wird und der Responder diese auch annimmt. Henrich u.a. (2004) haben in fünfzehn verschiedenen Kulturen diese Entscheidungssituation spielen lassen, die Abgaben bezüglich der dem Spieler zur Verfügung stehenden Summe variieren zwischen 15 Prozent und 50 Prozent (vgl. Abbildung 1).

Abbildung 1: Ultimatumspiel in fünfzehn indigenen Völkern (Quelle: Marlowe 2004, S. 177)



Auf die Erklärung für diese unterschiedlichen Abgaben kann an dieser Stelle nicht genau eingegangen werden. Die Autoren vermuten, dass der kulturelle und

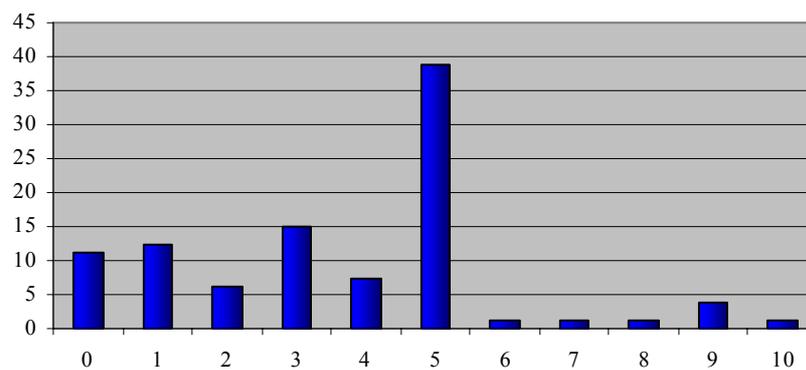
soziale Kontext ausschlaggebend ist: Das Ausmaß der Integration in die jeweilige Gesellschaft scheint unterschiedliche Ausprägungen von Normen zu begünstigen. So sind Kulturen, die im engen Familienverbund leben, tendenziell weniger fair im Diktatorspiel (siehe zum Beispiel Machiguenga). Kooperation mit Personen außerhalb des Familienkreises ist hier keine häufige Verhaltensweise. In westlichen Industriegesellschaften sind die Abgaben relativ ähnlich, es kommt keine solche Varianz vor wie in den Experimenten von Henrich et al. (2004). Aber auch hier scheinen Normen unterschiedlich ausgeprägt zu sein.

In diesem Beitrag soll am Beispiel von Experimenten an der RWTH Aachen veranschaulicht werden, inwieweit unterschiedliche Normen das Verhalten in Spielsituationen erklären können. Dafür wurde auf das einfache Diktatorspiel und auf das sequenzielle Diktatorspiel zurückgegriffen. Im einfachen Diktatorspiel bestimmt der Diktator (oder Proposer) die Aufteilung einer fixen Geldsumme zwischen sich und einem anonymen Spielpartner. Die standardtheoretische Prognose für rationale Egoisten ist, dass der Diktator alles für sich behält und seinem anonymen Spielpartner nichts oder nur wenig mehr als einen Nullbetrag zuweist. Das sequenzielle Diktatorspiel ist ein Diktatorspiel, das über zwei Runden gespielt wird, wobei die Spieler in der zweiten Runde die Rollen tauschen. In Runde 1 kann wie im einfachen Diktatorspiel der erste Spieler einen erhaltenen Geldbetrag beliebig aufteilen. In der zweiten Runde wird der gleiche Betrag von Spieler 2 aufgeteilt. Rückwärtsinduktion führt zum teilspielperfekten Gleichgewicht. In der zweiten Runde wird Spieler 2 alles für sich behalten und einen Betrag von null abgeben. In der ersten Runde sollte Spieler 1 alles für sich behalten und nichts abgeben, so die Prognose.

2. Altruismus und Fairness im einfachen Diktatorspiel

In einem ersten Experiment im Jahr 2005 konnten die Studierenden einen Betrag von 10 Euro aufteilen. Die Ergebnisse des einfachen Diktatorspiels zeigen, dass sich ein großer Teil der Untersuchungseinheiten fair verhält und die Hälfte des Betrages (5 Euro) an den anonymen Spielpartner abgibt. Aus Abbildung 2 ist ersichtlich, dass auch egoistische Motive eine Rolle spielen, da einige Personen den gesamten Betrag für sich behalten. Und es gab auch einige Personen, die wesentlich mehr als die Hälfte der Beträge an die Spielpartner abgaben.

Abbildung 2: Ergebnisse des einfachen Diktatorspiels an der RWTH Aachen 2005. Relative Häufigkeit der Beiträge für ALTER (n=80)

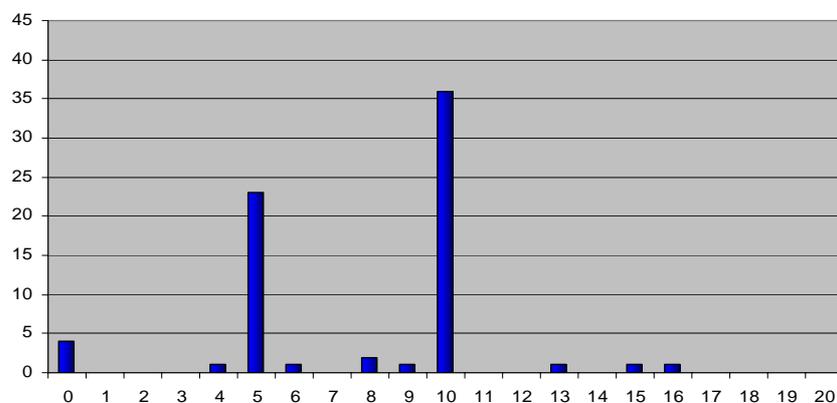


Das Experiment wurde ein Jahr später wiederholt, diesmal erhielten die Versuchspersonen 20 Euro. In der Aufteilung der Beträge ergibt sich eine ähnliche Verteilung wie im ersten Experiment (vgl. Abbildung 3). Damit hat die Höhe der Beträge offensichtlich keinen nennenswerten Einfluss auf die Entscheidung für die abgegebenen Beträge. Wiederum haben die meisten

Versuchspersonen die Hälfte der Summe an den Spielpartner abgeben. Egoistisch-rationale Entscheidungen treten selten auf.

Wie ist dieses Ergebnis einzuordnen? Zunächst stellt die Abgabe jeglicher Summe im Diktatorspiel eine altruistische Handlungsweise dar – wie oben genannt, ist es nicht rational dem anonymen Spielpartner ohne Schatten der Zukunft einen Betrag zu geben, zumal auch für den Spielpartner der Diktator anonym bleibt. Das Verhalten der Untersuchungseinheiten kann noch genauer spezifiziert werden: Die Spieler orientieren sich an einer paritätischen Aufteilung der Summe. Diese Präferenz für eine Gleichverteilung der Beträge kann auch als Fairness bezeichnet werden und ist nicht nur in unseren Experimenten in Aachen aufgetreten, sondern ist kennzeichnend für alle westlichen Industriegesellschaften.

Abbildung 3: Ergebnisse des einfachen Diktatorspiels an der RWTH Aachen 2006. Relative Häufigkeit der Beiträge für ALTER (n=71)



3. Ergebnisse des sequenziellen Diktatorspiels

Die Ergebnisse bislang können als Nachweis für altruistische oder auch faire Verhaltensweisen dienen. Mit dem sequenziellen Diktatorspiel soll nun auch der Nachweis für die Reziprozitätsnorm gelingen. Die Vorstellung, dass Reziprozitätsnormen weit verbreitet sind, wurde in der sozialwissenschaftlichen Literatur u.a. schon von Georg Simmel (1908) oder Alvin Gouldner (1960) postuliert. Reziprozität findet man als Strategie auch im iterierten Gefangenendilemma, bekannt unter dem Namen „Tit for Tat“ (Axelrod 2000). Diese Art von Reziprozität kann man im Unterschied zur Norm aber als strategische oder egoistische Reziprozität bezeichnen. Hier stellt freundlich-kooperatives Verhalten die beste Strategie rationaler Spieler dar und erhöht die Auszahlungen.

An dieser Stelle ist die altruistische Reziprozität als das Zurückgeben von Einsätzen oder Beträgen selbst in Entscheidungen, in denen die Spieler keine Zukunft mehr haben, von Interesse. Diekmann (2004) hat diese Art von Reziprozität mit dem sequenziellen Diktatorspiel nachgewiesen. Bei Diekmann wurden 3 Versuchsbedingungen konstruiert, die Versuchspersonen erhielten vom „ersten Spieler“ einen Anteil von 0,2, 0,5 oder 0,6 der 10 Punkte, (also 2, 5 oder 6 Punkte), die jeweils 2 Schweizer Franken Wert waren. Insgesamt sollte eine Summe von 20 Franken (entspricht etwa 14 Euro) aufgeteilt werden. Der erste Spieler ist in der Experimentsituation nicht real sondern fiktiv, jedoch ist dies den Versuchspersonen nicht bekannt. Etwa die Hälfte der Versuchspersonen geben genau den Anteil in der zweiten Runde zurück, den sie in der ersten Runde vom fingierten Spieler erhalten haben. Haben die Spieler also einen Anteil von 20% von der fingierten ersten Person erhalten, dann gaben die Versuchspersonen mehrheitlich auch 20% ihrer Summe ab. Haben sie 50% vom ersten Spieler erhalten, wurden von über 45% der Personen auch 50% zurückgegeben und erhielten sie 60%, dann teilten sogar an die 60% der Untersuchungseinheiten

ebenfalls 60% des Betrages aus. Diekmann interpretiert dies als Beleg für die Existenz einer Norm der Reziprozität. Die Ergebnisse können tatsächlich nicht durch reinen Altruismus erklärt werden, denn dann hätten in jeder Versuchsbedingung alles oder wenigstens gleich große Beträge abgegeben werden müssen.

Das sequenzielle Diktatorspiel wird in unseren Experimenten ebenfalls mit unterschiedlichen Versuchsbedingungen gespielt, allerdings mit einer Variation. In einer Bedingung erhalten die Spieler von ihrem Spielpartner (der erste Spieler ist wiederum fingiert) niedrige Vorleistungen, in der zweiten Bedingung bekommen sie hohe Vorleistungen. In jeder dieser Bedingungen gibt es Versuchspersonen, die die erhaltene Summe entweder an die fingierte Person und damit an ihren direkten Spielpartner zurückgeben oder an eine andere Person.¹ Bei der Möglichkeit der Zurückgabe eines Betrags an die Person, die einem in der Runde zuvor den Betrag ausgeteilt hat, kann spezifische (oder auch direkte) Reziprozität gemessen werden. Wenn die Aufteilung der Beträge nicht mehr an die spezifische Person möglich ist, sondern an eine andere Person, die mit der Aufteilung des ersten Betrags nicht zu tun hat, dann wird generalisierte (oder indirekte) Reziprozität gemessen. Spezifische Reziprozität meint damit, dass eine Vorleistung einer Person A an Person B danach mit einer Leistung der Person B an Person A zurückgezahlt wird. Generalisierte Reziprozität verlangt dagegen eine Transformationsleistung von Person B: Hierbei wird nicht ein direkter Tausch vorgenommen, sondern Person B erhält eine Summe (von 0 bis 20 in unserem Fall) von Person A und gibt an eine Person C weiter. Die Orientierung von Person B an Person A in der Höhe des abgegebenen Betrags

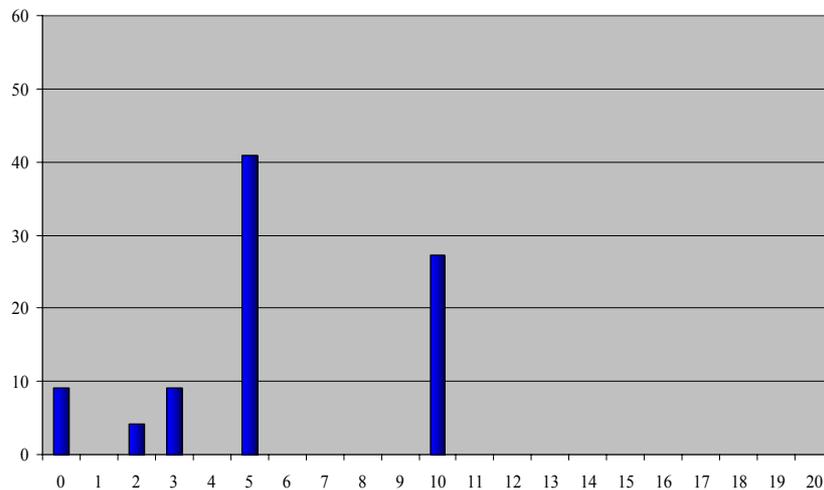
¹ Den Versuchspersonen wurde gesagt, dass diese Person zufällig ermittelt wird. Die Experimentalsituation wurde den Versuchspersonen so dargestellt, dass sie den Eindruck hatten, der erste Teil des Experiments hätte an der LMU München stattgefunden. Daher nahmen die Versuchspersonen an, dass die direkten Spielpartner (direkte Reziprozität) und die zufällig ermittelten Spielpartner (indirekte Reziprozität) Studierende an der LMU München sind.

bezeichnet das Ausmaß an generalisierter Reziprozität. Diese Versuchsbedingung der generalisierten Reziprozität hat natürlich auch Entsprechungen in realen Verhaltenssituationen. So leisten Menschen einen Obulus in öffentlichen Toiletten, bringen ihr Geschirr in Selbstbedienungsrestaurants zurück oder spenden Blut – in allen Fällen tritt man in Vorleistung (entweder in Form von Geldgaben oder Handlungen), obwohl zumindest keine direkte Rückleistung zu erwarten ist.

3.1 Spezifische negative und positive Reziprozität

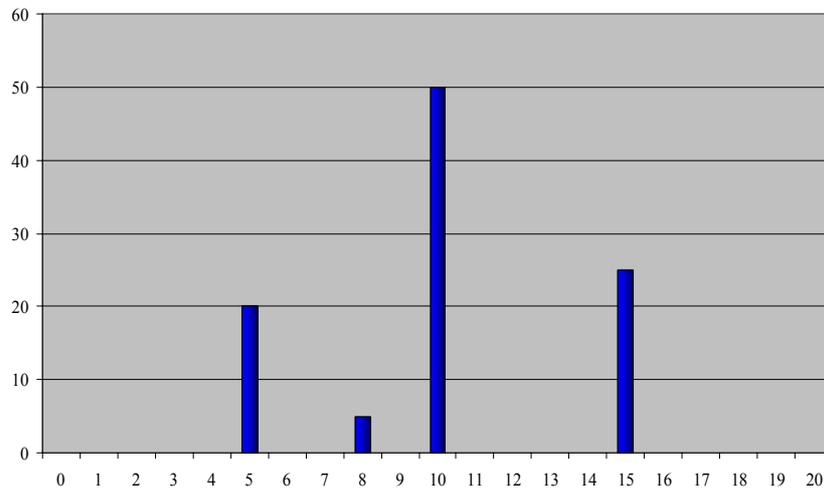
In der folgenden Abbildung 4 ist das Ergebnis der Versuchsbedingung mit geringen Vorleistungen und direkter Erwidern auf den fingierten Spieler zu sehen. In diesem Fall wird negative spezifische Reziprozität getestet: Für die Versuchspersonen war ersichtlich, dass der fingierte Spieler 20 Euro bekam und davon nur einen kleinen Teil (5 Euro) abgab. Erwartet wird, falls negative Reziprozität existiert, dass die Versuchspersonen von ihrem Betrag (20 Euro) ebenfalls nur geringe Beträge abgeben. Dies ist der Fall: Die meisten der Versuchspersonen antworten mit geringen Beträgen (Häufung bei 5 Euro). Relativ viele Spieler geben aber auch die Hälfte der Summe ab. Die Spieler scheinen trotz geringer Vorleistung eine Präferenz für *überfaire* Aufteilungen zu haben. Eine genau egalitäre Aufteilung in diesem Fall wäre eigentlich die Abgabe der Summe, die man erhalten hat. Damit hätte jeder Spieler absolut gleich viel erhalten. Die negative spezifische Reziprozität kann hier zwar festgestellt werden, aber der abgegebene Mittelwert von $\bar{x} = 5,65$ zeigt, dass tendenziell mehr gegeben wird, als unter Bedingungen der negativen Reziprozität zu erwarten ist.

Abbildung 4: Spezifische negative Reziprozität (gegeben wurden 5 Euro).
Relative Häufigkeit der Beiträge für ALTER (n=20)



Die nächste Versuchsbedingung testet hohe Vorleistung und deren Beantwortung, damit soll spezifische positive Reziprozität untersucht werden. Die Spieler erhalten von der fingierten Person 15 Euro. Positive Reziprozität würde vorliegen, wenn in etwa diese Beträge von der Versuchsperson abgegeben werden. Die Ergebnisse (Abbildung 5) zeigen, dass sich positive Reziprozität nur zu einem gewissen Ausmaß feststellen lässt (immerhin 25% der Personen). Die Spieler orientieren sich jedoch hauptsächlich an der Mitte der Verteilung, 50% der Untersuchungseinheiten geben 10 Euro ab, der Mittelwert liegt bei $\bar{x}=10,15$.

Abbildung 5: Spezifische positive Reziprozität (gegeben wurden 15 Euro).
Relative Häufigkeit der Beiträge für ALTER (n=20)



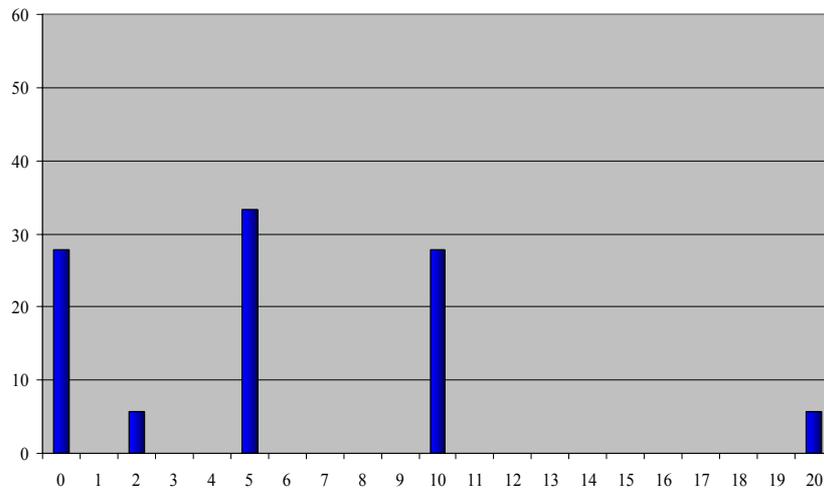
3.2 Generalisierte negative und positive Reziprozität

Im Folgenden werden Ergebnisse des Experiments dargestellt, in denen die Abgabe der Summe durch die Versuchsperson nicht mehr direkt an Spieler 1 erfolgt, sondern an einen anderen Spieler. Für die Versuchspersonen war aufgrund des Experimentbogens ersichtlich, dass der Spieler 1, der ihnen die Summe zugeteilt hat, auf keinen Fall etwas von ihrem Betrag bekommen kann.

In Abbildung 6 wird zunächst die Versuchsbedingung mit geringer Vorleistung betrachtet. Die Spieler haben 5 Euro von dem ersten Spieler bekommen und mussten entscheiden, wie viel von ihrem Betrag sie an einen ihnen unbekanntem Spieler abgeben wollen. Deutlich wird, dass unter den Bedingungen geringer Vorleistung auch zu einem großen Ausmaß geringe

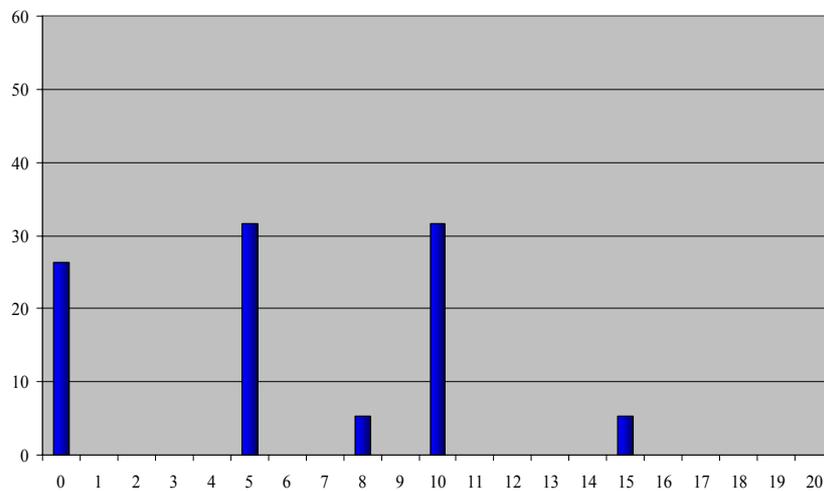
Beträge weitergegeben werden. 33,3% der Personen zeigen damit generalisierte negative Reziprozität und geben damit ihre Erfahrung geringer Auszahlung weiter. Ein Teil bestraft sogar in hohem Ausmaß, etwa 28% der Versuchspersonen geben nichts von ihrem Betrag ab. Allerdings gibt es auch einen ähnlich großen Anteil an Personen, welche die geringe Vorleistung nicht weitergeben, sondern fair agieren und die Hälfte der Summe abgeben. Wesentliches Ergebnis ist aber, das hier der Mittelwert von $\bar{x} = 5,7$ fast identisch ist mit dem Mittelwert der spezifischen negativen Reziprozität ($\bar{x} = 5,65$), d.h. die Versuchspersonen haben nicht zwischen diesen unterschiedlichen Bedingungen differenziert! Bei niedriger Vorleistung wird wenig weitergegeben, egal ob dies direkt an die erste Person ausgezahlt wird oder ob ein unbeteiligter Dritter der anonyme Spielpartner ist.

Abbildung 6: Generalisierte negative Reziprozität (gegeben wurden 5 Euro).
Relative Häufigkeit der Beiträge für ALTER (n=18)



Im Fall der generalisierten positiven Reziprozität wird aus Abbildung 7 deutlich, dass diese Form der Reziprozität nicht auftritt. Der Mittelwert der Abgaben liegt bei $\bar{x} = 5,95$ und ist damit kaum höher, als unter Bedingungen geringer Vorleistung. Die Verteilung besitzt zwei Modi, bei der Abgabe von 5 Euro und bei 10 Euro. Generalisierte positive Reziprozität, in diesem Fall die Abgabe von 15 Euro, tritt praktisch nicht auf.

Abbildung 7: Generalisierte positive Reziprozität (gegeben wurden 15 Euro).
Relative Häufigkeit der Beiträge für ALTER (n=19)



4. Zusammenfassung

Das Experiment demonstriert die Existenz spezifischer und generalisierter negativer Reziprozität: Spieler, die weniger als die Hälfte bekommen haben, geben auch nur wenig zurück. Merkwürdigerweise gilt das auch für das sequenzielle Diktatorspiel bei wechselnden Spielerpartnern.

Spezifische positive Reziprozität lässt sich nur eingeschränkt feststellen, die Versuchspersonen geben mehrheitlich nur etwa die Hälfte zurück, deutlich mehr jedoch als bei Bedingungen der spezifischen negativen Reziprozität. Die Ergebnisse zur generalisierten positiven Reziprozität zeigen, dass sich diese nicht bestätigen lässt. Bei hoher Vorleistung von Spieler 1, geben die Versuchspersonen dennoch nicht mehr als die Hälfte an einen dritten Spieler weiter. Vielmehr tritt eine große Varianz in den Abgaben zu Tage. Damit existiert weder spezifische positive noch generalisierte positive Reziprozität.

Ingesamt sind die Ergebnisse mit dem ERC Modell von Bolton und Ockenfels (2000) $u_i = a_i y_i - b_i (1/2 - \sigma_i)^2$ nicht kompatibel. Abweichungen von der Gleichverteilung gehen in dem Modell in jedem Fall negativ in die Nutzenfunktion ein (symmetrisch). Die experimentelle Befunde unterstellen aber eine asymmetrische Beziehung: $u_i = a_i y_i - b_i (1/2 - \sigma_i)$.

Literatur

- Axelrod, Robert (2000) [1984]: Die Evolution der Kooperation. 5. Aufl., München: Oldenbourg.
- Bolton, Gary E./Ockenfels, Axel (2000): ERC: A theory of equity, reciprocity, and competition. In: The American Economic Review 90: 166-193.
- Diekmann, Andreas (2004): The Power of Reciprocity. Fairness, Reciprocity, and Stakes in variants of the Dictator Game. In: Journal of Conflict Resolution 48: 487-505.
- Fehr, Ernst/Gächter, Simon (2000): Fairness and Retaliation: The Economics of Reciprocity. In: The Journal of Economic Perspectives 14: 159-181.
- Forsythe, Robert/Horowitz, Joel L./Savin, N.E./Sefton, Martin (1994): Fairness in simple bargaining experiments. In: Games and Economic Behavior 6: 347-369.

- Gouldner, Alvin W. (1960): The Norm of Reciprocity: A Preliminary Statement. *American Sociological Review* 25: 161-178.
- Henrich, Joseph/Boyd, Robert/Bowles, Samuel/Camerer, Colin/Fehr, Ernst/Gintis, Herbert (2004): *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.
- Ockenfels, Axel (1999): *Fairneß, Reziprozität und Eigennutz. Ökonomische Theorie und experimentelle Evidenz*. Tübingen: Mohr Siebeck.
- Simmel, Georg (1908): *Soziologie. Untersuchung über die Formen der Vergesellschaftung*. Frankfurt am Main: Suhrkamp.

Axel, Franzen, Prof. Dr., Lehrstuhl für Mikrosoziologie an der RWTH Aachen, Institut für Soziologie, Veröffentlichung u.a.: *Social Networks and Labour Market Outcomes: The Non-Monetary Benefits of Social Capital*. (mit Dominik Hangartner), erschienen in: *European Sociological Review* 22, 2006, pp. 353-368. Arbeitsschwerpunkte: Arbeitsmarktsoziologie, Bildungssoziologie, Spieltheorie, Sozialkapital, Sozialstruktur, Umweltsoziologie.

Pointner, Sonja, Dipl.-Soz., Institut für Soziologie der RWTH Aachen. Veröffentlichung: *Mobilität im Arbeitsmarkt* (mit Thomas Hinz), erschienen in: Martin Abraham/Thomas Hinz (Hrsg.): *Arbeitsmarktsoziologie: Probleme, Theorien, Empirische Befunde*. Wiesbaden: VS Verlag (2005). Arbeitsschwerpunkte: Arbeits- und Organisationssoziologie, Wirtschaftssoziologie, Sozialkapital.