

Kommentar zum Vortrag von Klaus M. Schmidt

Gerd Gigerenzer und Nathan Berg

Max-Planck-Institut für Bildungsforschung

Experimentelle Spiele bestechen durch die Klarheit der Regeln und die Einfachheit der Situation. Klarheit und Einfachheit kann helfen, komplexe Fragen einer empirischen Untersuchung zugänglich zu machen. Solche Untersuchungen haben die Grenzen der klassischen Spieltheorie für das Verständnis menschlichen Verhaltens aufgezeigt. In dieser Tradition, berichten Fehr, Klein und Schmidt (2007) über experimentelle Spiele zur Wahl zwischen Verträgen. Sie definieren drei Regeln für die Interaktion zwischen Arbeitgeber und Arbeitnehmer, die sie als „bonus contract“, „incentive contract“ und „trust contract“ bezeichnen und berichten Ergebnisse, die der Standard-Vertrags-Theorie widersprechen.

Wie können wir das abweichende Verhalten erklären? Fehr und Schmidt bieten zwei Erklärungen an. Die erste ist formal und wird als „inequality aversion“ Theorie bezeichnet. Die zweite ist verbal und besteht aus der Aussage, dass Menschen ein „concern for fairness and reciprocity“, also ein Bedürfnis nach Fairness und Reziprozität hätten.

Wir argumentieren:

1. Die Inequality-Aversion Theorie von Fehr und Schmidt bietet für das Verhalten in experimentellen Spielen wie dem Ultimatum Spiel nicht wirkliche Erklärung an, sondern ist weitgehend eine Wiederbeschreibung des Phänomens (d.h., die gleiche Verteilung des Gewinns).
2. Wenn es sich wie in der Arbeitgeber-Arbeitnehmer Situation dagegen nicht um Spiele handelt, in denen „windfall money“ verteilt wird, ist das unterstellte Motiv einer Gleichverteilung des Gewinns (inequality aversion) äußerst zweifelhaft.
3. Obgleich die Inequality-Aversion Theorie als „theories of fairness“ (Fehr et al., p. 121) dargestellt wird, bleibt unklar, was die formale und die verbale Erklärung miteinander zu tun haben.
4. Fehr et al. versuchen, Abweichungen von der klassischen Theorie des erwarteten Nutzens damit aufzufangen, indem man unerwartete Einflussgrößen als freie Parameter in die Nutzensfunktion mit hineinpackt. Dieses „Reparaturprogramm“ schützt die Theorie des erwarteten Nutzens zwar, aber unser Verständnis des menschlichen Verhaltens in den experimentellen Spielen wird dadurch nicht notwendig vergrößert. Eine Alternative ist der Ansatz der begrenzten Rationalität, welcher formale Modelle von Heuristiken entwickelt und Erklärungen anbietet (Gigerenzer & Selten, 2001).

I. Der theoretische Status der Inequality-Aversion Theorie.

Was besagt diese Theorie? Im Fall von zwei Spielern i und j ist der Nutzen $U_i(x)$ gegeben durch:

$$U_i(x) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}$$

In Worten, der Nutzen des Payoff-Vektors $x = (x_1, x_2)$ für Person i ist ihr eigener Geldbetrag x_i minus der gewichteten Differenz zwischen dem Betrag den sie selbst und der

andere Spieler erhält. Ist die Differenz null, so wird kein Nutzen abgezogen, ist die Differenz negativ, das heißt, Person j erhält mehr, so wird die Differenz mit α_j gewichtet, ist sie positiv, dann mit β_j . Gegeben ein Gewinn x_i , ist also der Nutzen für eine Person am größten, wenn beide Spieler gleich viel Geld erhalten.

Was sagt uns diese Theorie nicht? Sie lässt die beiden individuellen Parameter offen, und gibt nicht an, wie diese unabhängig von den Daten gemessen werden könnten. Das heißt, sie enthält keine theoretische Spezifikation der Parameter im Sinne einer Theorie von Fairness.

Wiederbeschreibung statt Erklärung. Betrachten wir zunächst experimentell Spiele wie das Ultimatum-Spiel, in denen der Experimentator einem Spieler eine Geldsumme gibt, die dieser nicht selbst durch Arbeit verdient hat („windfall money“), und letzterer dann einen Teil an den anderen Spieler weitergeben kann. Das bei vielen Spielern beobachtete Verhalten, das Geld gleich zu verteilen, wird nun dadurch „erklärt“, dass diese eine Inequality-Aversion hätten. Dabei handelt es sich also um eine Wiederbeschreibung des zu erklärenden Phänomens. Erinnern Sie sich an Molière’s Parodie der Aristotelischen Doktrin der Formen: Warum macht Opium schläfrig? Wegen seiner dormativen Eigenschaften. Dieses Problem könnte vermieden werden, wenn die zwei freien Parameter pro Person (α und β) unabhängig messbar wären und tatsächlich auch vor dem Experiment gemessen worden wären. Dann könnte man eine Vorhersage des Verhaltens machen, statt im nachhinein die beobachteten Daten mit zwei adjustierbaren Parametern pro Person (!) zu fitten. Da die Konzeption von Fehr et al. aber einer Psychologie der festen Persönlichkeitszüge entspricht, und wir für solche generelle Fairness-Traits wenig Evidenz haben, wird aber ein Versuch in dieser Richtung u. E. wahrscheinlich erfolglos bleiben.

II. Ist Inequality-Aversion relevant zum Verstehen der Arbeitsbeziehung zwischen Firma und Angestellten?

Anders als im Ultimatum Spiel (und ähnlichen Spielen), wo eine Person vom Experimentator Geld bekommt, das nicht selbst verdient worden ist („windfall money“), haben Arbeitgeber in der Regel das Geld verdient, das sie zu verteilen haben, und Gleichverteilung ist nicht ihre Motivation. Genauso vergleichen Arbeitnehmer ihren Verdienst in der Regel nicht mit dem des Arbeitgebers, sondern mit denen anderer Arbeitnehmer. Selbstverständlich beklagen sich oft Arbeitnehmer, dass der Profit für den Arbeitgeber schneller wächst. Die intuitive Logik eines Modells in dem der Nutzen für beide in kritischer Weise von der Gleichheit beider Profite abhängt, ist jedoch schwer nachzuvollziehen. Genauso hat Inequality-Aversion seine Grenzen, wenn der „Proposer“ im Ultimatumspiel sein Geld erst einmal verdienen musste, anders als mit „windfall money.“

In Bezug auf die Asymmetrie zwischen Arbeitgeber und Arbeitnehmer besagt eine der Vorhersagen der Theorie von Fehr et al., dass jene Partei in den Vertragsverhandlungen, die weniger zu verlieren hat, auch eher bereit sein sollte, zu vertrauen. So argumentieren sie, dass in der wirklichen Welt die Bonusverträge den Vertrauensverträgen vorgezogen werden, da Arbeitnehmer weniger als Arbeitgeber zu verlieren hätten und daher eher geneigt seien, dem Arbeitgeber zu vertrauen, dass er den Bonus zahlt. Dagegen hätte der Arbeitnehmer weniger Vertrauen darin, dass der Arbeitnehmer eine hohe Anstrengung leisten wird. Fehr et al. haben sicher recht, dass es in der realen Welt wichtige Asymmetrien bezüglich politischer Macht, Verhandlungsposition und finanzieller Mittel gibt. Aber ohne entsprechende Daten erscheint es

uns wenig begründet, anzunehmen, dass Arbeitnehmer weniger zu verlieren haben – man denke nur an Wohnung, medizinische Versorgung, Arbeitslosigkeit, usw. als Folge eines schlechten Ergebnisses des Arbeitsmarktes.

III. Ist „Inequality-Aversion“ das gleiche wie „Fairness Concern“?

Fehr et al. (2007) behandeln „inequality aversion“ und „concern for fairness“ oder „concern for reciprocity“ als austauschbare Erklärungen. Eine gleiche Aufteilung von Geld ist jedoch ein Spezialfall einer fairen Aufteilung und nicht gleichzusetzen mit Fairness. Wie vorher argumentiert, kann diese Definition eine faire Aufteilung in „windfall“ Situationen charakterisieren, nicht aber unbedingt bei Verträgen zwischen Firma und Arbeitnehmer. Wahrgenommene Fairness reagiert sensitiv auf vielen wichtige Randbedingungen, und das ist aus moralischer Sicht auch so verständlich. Die Inequality-Aversion Theorie befasst sich mit keiner dieser Randbedingungen. Vielmehr ist sie durch die beiden adjustierbaren Parameter pro Person in hohem Masse unterspezifiziert, was gleichfalls für die verbale Umschreibung „concern for fairness“ gilt.

IV. Heuristische Prozesse statt Maximierung mit Trait-Psychologie

Das Programm von Fehr und Schmidt nimmt die Diskrepanz zwischen der Theorie der Maximierung des erwarteten Nutzens (EUT) und der experimentellen Evidenz ernst (was nicht alle Ökonomen tun). Wir begrüßen dies. Statt aber eine Theorie zu entwickeln, welche die Entscheidungsprozesse modelliert, halten sie an der EUT fest, und erweitern die Nutzensfunktion um zwei individuelle Parameter, welche „inequality avoidance“ als Teil des Nutzens repräsentieren sollen. Diese situations-unabhängigen Parameter entsprechen dem was Psychologen als stabile Persönlichkeitszüge (Traits) bezeichnen, und wofür im gegebenen Fall die empirische Evidenz sehr fraglich ist. Wir sind der Meinung, dass dieser Weg in eine theoretische Sackgasse führt. Falls Fehr et al. die intraindividuelle Konsistenz der Fairness-Parameter über verschieden Situations überprüfen würden, dann werden sie mit an Sicherheit grenzender Wahrscheinlichkeit finden, dass diese gering ist, konsistent mit der psychologischen Literatur.

Es gibt eine theoretische Alternative zu dem Fehr et al. Programm, abweichendes Verhalten als Disposition in die Nutzensfunktion zu packen. Dies sind formale Modelle von Heuristiken welche keine Berechnungen von Nutzensfunktionen enthalten, sondern versuchen, die Entscheidungsprozesse direkt zu modellieren (e.g., Gigerenzer et al., 1999). Die adaptive Verwendung von Heuristiken erlaubt situations-spezifisches Verhalten. Tit-for-tat ist ein klassisches Beispiel, das zeigt, wie Verhalten nicht mittels Wiederbeschreibung mittels einer Disposition „erklärt“ wird. Die gleiche Heuristik kann zu kooperativen oder aber zu nicht-kooperativen Verhalten führen, je nach dem Verhalten des anderen Spielers. Es wäre verfehlt, von dem beobachteten Verhalten auf eine zugrunde liegende Disposition der Kooperation oder Kooperations-Aversion zu schließen. Mit Dispositions-Parametern kann man nicht adaptives Verhalten erfassen.

Wie können Heuristiken die Präferenz einer Vertragsform über eine andere erklären? Ein Vorschlag wäre, sequentielle Heuristiken wie Take The Best zu testen (Gigerenzer et al., 1999). Diese haben die Form: Bei der Wahl zwischen zwei Verträgen, vergleiche diese auf

Eigenschaft A. Falls sie sich hier unterscheiden, nimm den Vertrag, der auf A den besseren Wert hat und ignoriere alle andere Information. Andernfalls, gehe zu Merkmal B, und so weiter. Viele Menschen treffen Entscheidungen in dieser sequentiellen Art und Weise. Eine mögliche Hypothese für die Eigenschaften im vorliegendem Fall wären Anzahl und Sicherheit der Dimensionen der Kontrolle: Wenn ich bei einem Vertrag mehr Dimensionen kontrollieren kann, dann nehme ich diesen. Ist die Anzahl jedoch gleich, dann nehme ich den Vertrag bei dem die Kontrolle sicherer ist. Eine solche Heuristik würde erklären, warum „incentive contracts“ den „trust contracts“ vorgezogen werden, da der erstere eine Kontrollvariable (die Bestrafung) mehr hat. Sie würde ebenfalls vorhersagen, dass „bonus contracts“ den „incentive contracts“ vorgezogen werden, da hier die Anzahl der Kontrollvariable gleich ist, aber der Bonusvertrag eine sichere Kontrollvariable hat (der Bonus), während der „incentive contract“ nur mit einer kleinen Wahrscheinlichkeit eine Kontrolle erlaubt. Diese Heuristik ist eine Hypothese, die wir hier vorschlagen, um nicht nur kritisch sondern konstruktiv zu dem Projekt von Fehr, Klein und Schmidt beizutragen.

In der Diskussion im Anschluss an seinen Vortrag hat Klaus Schmidt betont, dass er keinen Widerspruch zwischen seinem Programm und unserem Heuristik-Programm sieht, sondern dass beides ähnliche Erklärungen seien. Wir sehen das anders: beide Ansätze unterscheiden sich in zentralen Punkten, und man muss zeigen, welches das Verhalten besser vorhersagt (nicht: die Daten im Nachhinein fittet) und auch erklärt. Zu diesen Punkten gehören: Sollen wir versuchen menschliches Verhalten durch einen einzigen Kalkulus, der Maximierung des erwarteten Nutzens, erklären? Und falls dieser dem beobachteten Verhalten widerspricht, sollen wir dann auch diese Abweichungen in die Nutzensfunktion hinein nehmen? Reinhard Selten hat dieses Vorgehen zur Rettung der EUT als „Reparaturprogramm“ bezeichnet. Oder sollen wir Verhalten als Funktion einer „adaptive toolbox“ von Heuristiken erklären, welche für definierte Problemklassen adaptiv sind (Gigerenzer & Selten, 2001)? Ist Maximierung ein psychologisch realistischer Erklärungsansatz obgleich die meisten realen Probleme keine Maximierung erlauben? Beispielsweise sind Bayesianische „belief networks“ NP-hart, das heißt, die optimale Lösung kann nicht gefunden werden (Cooper, 1990) und das gilt auch für Approximierungen. Selbst in der von Klarheit und Einfachheit geprägten Welt eines Spiels in dem die Wahl auf nur drei Vertragsarten begrenzt ist, muss man darüber nachdenken, ob die auf Maximierung beruhenden Ergebnisse auf die institutionellen und strategischen Möglichkeiten der wirklichen Welt verallgemeinerbar sind. Diese Grenzen der Maximierung illustrieren, dass in vielen komplexen Umwelten heuristische Prozesse unverzichtbar sind, und Maximierung zur Science Fiction wird.

- Cooper, G. (1990). The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence*, 42, 393–405.
- Fehr, E. Klein, A. & Schmidt, K. M. (2007). Fairness and contract design. *Econometrica*, 75, 121-154.
- Gigerenzer, G., & Selten, R. (Eds.). (2001). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.