

**MARKTDESIGN UND EXPERIMENTELLE
WIRTSCHAFTSFORSCHUNG**

AXEL OCKENFELS

Marktdesign und Experimentelle Wirtschaftsforschung

Axel Ockenfels*, Universität zu Köln

22. November 2008

1. Einleitung: Marktdesign als angewandte Forschungsdisziplin

Marktdesign ist die Kunst, Institutionen so auszugestalten, dass Anreize und Verhalten mit den übergeordneten Zielen im Einklang stehen. Die Theorie des Marktdesigns definiert Institutionen als nicht-kooperative Spiele und unterstellt rationales Verhalten. Auf Basis der jeweiligen Gleichgewichtsergebnisse können dann die Vor- und Nachteile von Institutionen miteinander verglichen werden.¹ Marktdesign als angewandte Wissenschaft beschäftigt sich komplementär mit der Detailfülle real existierender Institutionen und den Eigenheiten realen Verhaltens. Theorie und Anwendung des Marktdesigns befruchten sich durch ihre Erkenntnisse und Herausforderungen gegenseitig.

Beispiele für Fragen, mit denen sich Marktdesigner beschäftigen, sind: Welche Marktregeln führen zu effizientem Emissionshandel, und welche Organisationsform der europäischen Strombörse generiert wettbewerbliche Strompreise? Wie können innerbetriebliche Anreizsysteme die Mitarbeiter bestmöglich motivieren? Welche Informationsflüsse erzeugen in anonymen Online- oder Finanzmärkten Vertrauen, und wie können Anreize geschaffen werden, dass die notwendigen Informationen unverzerrt bereitgestellt werden? Nach welchen Zuordnungsregeln sollten Schüler auf Schulen, Ärzte auf Krankenhäuser und Organspender auf Organempfänger verteilt werden? Durch welche Auktionsregeln werden bei Finanzmarktaktionen, Privatisierungen oder im betrieblichen Einkauf Erlöse beziehungsweise Effizienz maximiert?

Die konzeptionellen Einsichten der Wirtschaftstheorie sind bei der Beantwortung solcher praxisrelevanten Fragen unentbehrlich. Zugleich reicht Theorie allein in der Regel nicht aus. Der Grund ist, dass die Modelle weder alle relevanten Details realer Märkte einfangen

* Mein Dank gilt der Deutschen Forschungsgemeinschaft für finanzielle Unterstützung, Achim Wambach für hilfreiche Kommentare sowie meinen Koautoren und Unterstützern in verschiedenen Marktdesign-Projekten, insbesondere auch Gary Bolton und Ben Greiner. Meine Einschätzungen in diesem Artikel werden nicht notwendigerweise von meinen Koautoren oder den beteiligten Unternehmen, Marktplattformen und Administrationen geteilt.

¹ Die Theorie des Mechanismusdesigns begann mit Hurwicz (1960, 1972); sie stand 2007 mit der Verleihung des Wirtschaftsnobelpreises an Leonid Hurwicz, Eric Maskin und Roger Myerson im Blickpunkt der Öffentlichkeit. Einführende Übersichtsartikel sind bei Baliga und Maskin (2003) und Serrano (2004) zu finden.

(Kapitel 2), noch allen relevanten Phänomenen realen menschlichen Verhaltens gerecht werden können (Kapitel 3). Angewandtes Marktdesign erfordert daher den komplementären Einsatz empirischer, experimenteller und theoretischer Methoden, die die relevanten Komplexitäten adressieren können. Dieser Artikel beschäftigt sich mit der Rolle der Experimentellen Wirtschaftsforschung bei angewandten Fragen des Marktdesigns (Kapitel 4).

Angewandtes Marktdesign ist gewissermaßen per Definition eine Forschungsdisziplin, die nicht vorrangig nach größtmöglicher Verallgemeinerung von Handlungsempfehlungen strebt, sondern sich mit dem ‚engineering‘ Aspekt konkreter Märkte beschäftigt (Roth 2002). Handwerkliche Fehler bei der Ausgestaltung der Mechanismen können dabei zu enttäuschenden oder gar katastrophalen Ergebnissen führen, wie Beispiele bei der Versteigerung von Frequenzbändern, beim Design von Energiemärkten oder bei der Fehlregulierung von Finanzmärkten nachdrücklich vor Augen führen. Ein vorrangiges Ziel des angewandten Marktdesigns ist daher, eine wissenschaftliche Literatur darüber aufzubauen, was konzeptionell für die Theorieentwicklung aus erfolgreichen und weniger erfolgreichen Designversuchen gelernt werden kann, und mit welchem Methodenmix trotz der Komplexitäten realer Institutionen und realen Verhaltens wissenschaftlich fundiert Hilfestellung gegeben werden kann. Kapitel 5 berichtet exemplarisch von einer aktuellen Studie, die insbesondere die Möglichkeiten der Experimentellen Wirtschaftsforschung beim angewandten Marktdesign illustriert. Kapitel 6 fasst zusammen.

2. Institutionelle Details

Eine Kernherausforderung des angewandten Marktdesigns ist die Untersuchung der Robustheit von Mechanismen. Finanzmärkte, Strommärkte oder Emissionshandelsmärkte, die unter mehr oder weniger extremen Bedingungen versagen, sind offenbar nicht robust. Zuweilen können bereits scheinbare Kleinigkeiten über den Erfolg oder Misserfolg von Mechanismen entscheiden. Zudem können dieselben Mechanismen in unterschiedlichen sozialen und ökonomischen Einbettungen zu ganz unterschiedlichen Ergebnissen führen (siehe zum Beispiel Klemperer 2004 für Illustrationen).

Wirtschaftstheoretische Modelle vermitteln eine hilfreiche Intuition darüber, wie Märkte und Anreizsysteme auf Verhalten und Marktergebnisse wirken können. Dafür abstrahieren sie oft von institutionellen Details, da andernfalls eine mathematische Analyse erschwert, die Allgemeingültigkeit der Schlussfolgerungen eingeschränkt oder die Intuition behindert wird. Angewandtes Marktdesign darf solche Komplexitäten nicht ausblenden. Ob und welche Details in Institutionen und im Verhalten relevant sind, können Robustheitschecks offenbaren.

Robustheitschecks sind – anders als bei anderen angewandten Wissenschaften – noch recht untypisch für die Wirtschaftswissenschaft.² Flugzeugbauingenieure sind zwar auch mit

² Es gibt natürlich auch Zweige in der Theorie, die sich mit der Robustheit ihrer Aussagen beschäftigen. Wilson (1985) schreibt zum Beispiel, dass es wünschenswert sei, dass ein Mechanismus “does not rely on features of the

theoretischen Grundlagenkenntnissen (u.a. in den Bereichen Aerodynamik, Thermodynamik, Strukturmechanik, Flugmechanik, Regelungstechnik, Antriebstechnik, Werkstofftechnik, Datenverarbeitung) ausgestattet. Doch sie werden in der Regel nicht darauf verzichten, die Robustheit eines neu zu entwickelnden Flugzeugsystems auch im Labor ausgiebig zu testen. Dabei werden Millionen elektronische und andere Komponenten unter den relevanten Umgebungsparametern auf Funktionsfähigkeit geprüft, und der Einfluss des Versagens der Einzelkomponenten oder von Paketen solcher Komponenten auf das Gesamtsystem untersucht (einschließlich menschlichem Versagen). Niemand wäre bereit, ein neues System ohne solche wohldefinierten Tests in die Praxis zu entlassen. In der Wirtschaftswissenschaft sieht das anders aus. Es gibt keine Standards in der ökonomischen Literatur für Robustheitschecks bei Märkten und Anreizsystemen. Auch gibt es nur wenig Hilfestellung, wie die Blaupausen der Wirtschaftstheorie handwerklich robust und mit Blick fürs Detail in funktionsfähige Mechanismen umgesetzt werden können. Angewandtes Marktdesign möchte diese Lücke füllen.

Im Folgenden sollen beispielhaft zwei Komplexitäten beschrieben werden, die durch die Wirtschaftstheorie tendenziell vernachlässigt werden, die sich jedoch beim angewandten Marktdesign zuweilen als kritisch herausgestellt haben (und auch beide noch in Kapitel 5 eine Rolle spielen werden): der Einfluss der Kommunikationskanäle und der endogenen Dynamik von Aktionen auf das Marktergebnis.³

2.1 Kommunikationskanäle

Die Ökonomik besitzt keine Theorie, die deskriptiv erfolgreich den weitgehenden Einfluss von Kommunikation und ihrer Kanäle bei Verhandlungskonflikten und unvollständigen Verträgen beschreiben oder prognostizieren könnte. Tatsächlich gilt nämlich, dass zur Überwindung solcher Probleme kein Mittel so erfolgreich zu sein scheint wie Kommunikation – und zwar auch dann, wenn sie theoretisch als cheap talk (dis)qualifiziert wird (Ledyard 1995, Sally 1995, Ockenfels und Selten 2000, Brosig et al. 2003, Brosig 2006).

Kommunikationskanäle geraten in der ökonomischen Praxis oft in den Mittelpunkt der Diskussionen. Ein prominentes Beispiel ist das Wettbewerbsrecht, das einen entscheidenden Unterschied zwischen impliziten und expliziten Preisabsprachen macht, obwohl dies aus wirtschaftstheoretischer Sicht jedenfalls nicht ohne weiteres gerechtfertigt werden kann (Whinston 2008). Analoge Phänomene gibt es in der Auktionsdesignliteratur. Offene Auktionen, in denen gleichzeitig mehrere Objekte versteigert werden, vereinfachen nämlich

agents' common knowledge, such as their probability assessments", und Dasgupta and Maskin (2000) "seek auction rules that are independent of the details – such as functional forms or distribution of signals – of any particular application and that work well in a broad range of circumstances."

³ Eine weitere Komplikation, die oft beschrieben wird, sind allokativen und informationellen Komplementaritäten beziehungsweise Externalitäten, die zu einem ganzen Bündel von institutionellen und verhaltenswissenschaftlichen Herausforderungen führen; siehe Milgrom (2004) sowie Jehiel und Moldovanu (2006) für Theorie, Brunner et al. (2007) für ein Experiment im Rahmen des Designs der ersten kombinatorischen Auktion der FCC sowie Ockenfels et al. (2006) zur Relevanz von Komplementaritäten im Online-Auktionssektor für angewandtes Marktdesign.

direkte Kommunikation zwischen den Bietern. Durch die Möglichkeit dynamisch-reziproker Bestrafung und Belohnung sowie impliziter koordinierender Absprachen über die einzelnen Runden hinweg erhöht die Offenheit die Gefahr von Kollusion. So ist es beispielsweise in Spektrumauctionen vorgekommen, dass Bieter mit den letzten Ziffern ihrer Gebote erfolgreich Absprachen und Nichtangriffspakte getroffen haben. Bereits die einfache Einschränkung, Gebote nur aus einem vorgegebenen Menü wählen zu dürfen, kann derlei Koordination und dadurch bestimmte Kollusionsstrategien wirkungsvoll verhindern (Klemperer 2004). Auktionsdesign muss eine Balance finden, in der zum einen die Kommunikationsmöglichkeiten der Bieter untereinander möglichst beschnitten werden, ohne aber den dynamisch-effizienten Preisfindungsprozess zu beeinträchtigen. Ein Weg dazu kann sein, nur aggregiertes und gegebenenfalls mit stochastischen Störungen versehenes Feedback über gebotene Mengen und Preise zu geben. Diese und ähnliche Überlegungen haben beispielsweise bei den jüngeren Auktionen virtueller Kraftwerkskapazitäten und von Strom in Deutschland eine Rolle gespielt.

Kapitel 5 gibt ein Beispiel, wie Kommunikationskanäle für Vertrauen sorgen und soziale Dilemma Probleme abschwächen können, und dadurch indirekt die Markteffizienz beeinflussen können (siehe auch Ostrom 1990).

2.2 Endogene Dynamik

Eine weitere Herausforderung ergibt sich durch die in realen Märkten oft existierende Flexibilität bei der Wahl des Zeitpunktes von ökonomischen Transaktionen. Spieltheorie unterstellt in der Regel vereinfachte, exogene Strukturen der Informationsflüsse – und abstrahiert so von der endogenen Dynamik der Aktionen in Echtzeit. Dass dieses Marktergebnisse signifikant beeinflussen kann, illustrieren folgende Beispiele.

Matching. In einer Reihe von Arbeiten zu angewandtem Marktdesign zeigt Alvin Roth, teils mit Koautoren, dass die Wahl des Zeitpunkts ökonomischer Aktionen eine entscheidende Rolle für das Funktionieren oder Versagen von Matching-Märkten spielen kann. Seine Arbeiten beschäftigen sich mit dem Design von Arbeitsmärkten für Ärzte und Ökonomen, dem Design von Prozeduren zur Organallokation bis hin zur Wahl von Schulen in New York City und Boston (siehe etwa Roth 1990, 1991, Mongell und Roth 1991, Roth und Xing 1994, Roth und Peranson 1999, Roth, Sönmez und Ünver 2004; Kagel und Roth 2000 sowie Chen und Sönmez 2006 sind Beispiele für komplementäre experimentelle Studien; Roth 2008 gibt einen Überblick). Eine der typischen Herausforderungen bei Matching-Märkten ist, dass aus Effizienzgründen zwar auf der einen Seite hinreichend viele Teilnehmer möglichst zur selben Zeit attrahiert werden müssen, aber zugleich auf der anderen Seite Überfüllungseffekte vermieden werden müssen. Dafür muss zum Beispiel der für viele solche Märkte typische Anreiz vermieden werden, *früher* als die Anderen handeln zu wollen. Roth zeigt, dass zentrale Zuteilungsmechanismen die resultierenden ineffizienten Dynamiken in den Griff bekommen können.

Sniping. In einer anderen Serie von Arbeiten, in denen das Zusammenspiel zwischen Theorie, Feld- und Laborforschung für ökonomisches Design besonders zentral ist, steht der Anreiz, *später* zu handeln als die Anderen, im Mittelpunkt des Interesses. Dabei geht es um die endogene Dynamik von Geboten und dem Design der Regel, nach der elektronische Auktionen enden. In Roth und Ockenfels (2002) wird beobachtet, dass sich Online-Auktionen zwar in den Preisregeln oftmals nicht unterscheiden, aber in der Regel, nach der die Auktionen enden. Insbesondere benutzen sowohl eBay als auch Auktionen, die von Amazon in den USA angeboten wurden, Zweitpreis-Regeln zur Preisbestimmung. Anders als bei der Vickrey-Auktion finden die Online-Auktionen jedoch in einem dynamischen, offenen Bietkontext statt. Amazon-Auktionen enden dabei mit einem *weichen Ende*, bei dem eine Auktion solange über die angestrebte Laufzeit hinaus verlängert wird bis 10 Minuten lang kein Gebot eingegangen ist. eBay-Auktionen enden dagegen zu einem bis auf die Sekunde vorgegebenen *festen Ende*. Gleichzeitig wird in einem natürlichen Experiment, bei dem Felddaten über die Handelsaktivitäten auf den beiden Auktionsplattformen untersucht werden, beobachtet, dass auf eBay ein signifikanter Teil der Gebote erst in den letzten Minuten beziehungsweise Sekunden abgegeben wird. In der Hektik der Schlussphase können die Preise steil ansteigen und Bieter bei dem Versuch scheitern, noch rechtzeitig auf späte Gebote zu reagieren. Auf Amazon ist dagegen ein solcher Effekt nicht zu beobachten – der Preisfindungsprozess ist gleichmäßiger über die Zeit verteilt, ohne allzu erhöhte Bietaktivitäten kurz vor dem geplanten Ende der Auktion. Der Unterschied zwischen eBay und Amazon wird größer, wenn die Güter mehr common-value-Charakteristiken aufweisen, und wenn die Bieter erfahrener sind.

In Ockenfels und Roth (2006) und Ariely, Ockenfels und Roth (2005) werden die Phänomene dynamischen Bietens theoretisch wie experimentell näher untersucht. In Ockenfels und Roth werden spieltheoretische Modelle von eBay und Amazon entwickelt und es wird gezeigt, dass spätes Bieten (*Sniping*) durch eine ganze Reihe von strategischen Gründen erklärt werden kann. *Sniping* kann insbesondere eine Gleichgewichtsstrategie in dynamischen Auktionsmodellen mit privaten oder interdependenten Bewertungen sein, ist aber auch außerhalb des Gleichgewichts eine vernünftige Strategie, denn es ist eine beste Antwort auf weit verbreitete (naive) inkrementelle Verhaltensmuster. Zugleich zeigt sich, dass die strategischen Anreize zum *Sniping* in Amazon-Auktionen wegfallen – während eine Reihe von psychologischen oder technischen Erklärungen der Bietdynamik keine Unterschiede prognostizieren würde. Felddaten können die Prognosen im Wesentlichen stützen.

In Ariely et al. (2005) werden eBay und Amazon im Labor nachgestellt, um unter kontrollierten Bedingungen Kausalität nachzuweisen. Das Experiment erlaubt zudem – anders als die Felddaten – zu untersuchen, in welcher Relation die individuellen Gebote zu den induzierten Bewertungen stehen, ob und auf welche Weise Lernen stattfindet, welche der konkurrierenden Theorien falsifiziert werden können und schließlich welche Effizienzimplikationen die verschiedenen Regeln besitzen. Ein Ergebnis ist, dass die dynamische Interaktion naiven *und* strategischen Verhaltens ein maßgeblicher Grund für die im Feld beobachteten Verhaltensmuster zu sein scheint. Zudem gilt, dass *ceteris paribus* das weiche Ende von Amazon den Preisfindungsprozess günstig beeinflusst und daher zu einer höheren Effizienz führt.

Insgesamt erzeugt die theoretische, empirische und experimentelle Literatur ein robustes Bild des Einflusses der Auktionsregeln auf das endogene Timing der Bietaktivitäten in Online-Auktionen.⁴ Analoge Herausforderungen für angewandtes Marktdesign existieren überall dort, wo strategische und dynamische Interaktion über elektronische Plattformen stattfindet, wie zum Beispiel auf diversen Verhandlungsplattformen (Ockenfels 2001) und in Börsen (siehe auch nächsten Abschnitt). Die Deutsche Börse lässt beispielsweise ihre Finanzauktionen innerhalb eines Zeitfensters zu einem zufälligen und unbekanntem Zeitpunkt enden, um die Anreize zum Sniping abzdämpfen. Im Ergebnis können auch bei diesem Design Gebote verloren gehen.

Strom. Strommärkte sind ein besonders aktives Feld der Marktdesignforschung (Wilson 2002, Müsgens und Ockenfels 2006), in dem auch Timing eine bedeutende Rolle spielt. Ein Grund ist, dass Angebot und Nachfrage zu jedem Zeitpunkt perfekt ausgeglichen sein müssen, da es sonst zu Instabilitäten im Netz und damit zu kostspieligen Zusammenbrüchen der Stromversorgung kommen kann. Zugleich unterliegen Angebot und Nachfrage starken und signifikanten Schwankungen, auch in Echtzeit, die etwa mit der Windeinspeisung und schwer zu prognostizierenden Ausfällen von Kraftwerken zu tun haben. Ein effizienter Strommarkt muss daher möglichst nahe an der Echtzeit geräumt werden, und bedarf zusätzlich der aktiven Unterstützung von Ausgleichsmechanismen und Regelenergiemärkten, die ihrerseits komplexes Marktdesign erfordern (Stoft 2002). Dabei sind die zeitlichen Restriktionen bindend, weil die Einsatzplanung der Kraftwerke vorbereitet werden muss und neuerdings zusätzlich auch die Kopplung europäischer Börsen zusätzliche wertvolle Zeit verbraucht.

Ein weiteres Problem ist, dass Strommärkte durch die dynamische Fahrweise der Kraftwerke mit diversen Komplexitäten kämpfen müssen, die etwa durch Komplementaritäten in den Kosten (beispielsweise Anfahr- und Abstellkosten von Kraftwerken) und Problemen der dynamischen Einsatzoptimierung (zum Beispiel bei Pumpspeichern) erzeugt werden. Dies hat Konsequenzen für das Auktionsdesign, und insbesondere für das Design der Gebotsformate. Können nämlich die Gebote die einzelnen Kostenbestandteile nicht hinreichend reflektieren und die individuelle Optimierung unterstützen, wäre Ineffizienz unausweichlich.

Theoretisch optimale Auktionsalgorithmen sind jedoch unpraktikabel. Erstens ist im allgemeinen Fall die Bestimmung der optimalen Allokation in kombinatorischen Problemen ein notorisch schwieriges Problem, das bei einer großen Anzahl von Bietern und potenziellen Kombinationen überhaupt nicht oder zumindest nicht in vernünftiger Zeit gelöst werden kann (siehe de Vries und Vohra 2003). Zweitens erfordert vollständige Effizienz nicht nur beliebig konditionale Gebotsformate, einschließlich sich gegenseitig ausschließender Gebote, sondern auch nichtlineare Preise, die ihrerseits die Komplexität der Preisfindung und des Bietens signifikant erhöhen und gleichermaßen die Transparenz des Verfahrens reduzieren. Praktisches Marktdesign konstruiert daher auf Basis der theoretischer Überlegungen

⁴ In dem Übersichtsartikel von Ockenfels, Reiley und Sadrieh (2006) wird die mittlerweile umfangreiche Literatur zur Dynamik von Bietaktivitäten präsentiert und eine Übersicht über die neuesten Entwicklungen in der ökonomischen Forschung zu Online-Auktionen gegeben; siehe auch Kittsteiner und Ockenfels (2006 und 2008) zu weiteren Designanstrengungen bei eBays Mehrgüterauktion.

praktikable und schnelle Suchalgorithmen; in Simulationsstudien wird dann ein Maß für die resultierende Ineffizienz bestimmt.

Marktdesign kann die Probleme der dynamischen Einsatzoptimierung von Kraftwerken auch im Detail abdämpfen helfen. Negative Strompreise, wie sie die EEX seit kurzem erlaubt, sind beispielsweise in einer zweitbesten Welt sinnvoll; nämlich dann, wenn es teurer ist, ein Kraftwerk in der nächsten Stunde herunterzufahren als es weiter laufen zu lassen. Wird den Anbietern nicht gestattet, negative Grenzkosten durch die Gebote auszudrücken, kann die Kosteneffizienz der Stromerzeugung leiden oder die Markträumung erschwert werden.⁵

Hilfreich kann zudem ein so genanntes ‚flexibles Stundengebot‘ sein. Dabei handelt es sich um ein Verkaufsgebot mit Reservationspreis und fester Menge, das automatisch der Stunde mit dem höchsten Markträumungspreis zugerechnet wird, falls der Markträumungspreis über dem Reservationspreis liegt. Dieses Gebotsformat erleichtert den Bietern das Lösen intertemporaler Optimierungsprobleme, da Preisprognosen für bestimmte Anwendungen überflüssig werden. Dies gilt einerseits für große industrielle Stromverbraucher auf der Nachfrageseite, die bereit sind, im Fall sehr hoher Marktpreise ihre Produktion für einen begrenzten Zeitraum auszusetzen, um den Strom aus ihren bestehenden Verträgen am Markt verkaufen zu können. Auf der Angebotsseite profitieren Kraftwerke, die in einer gewissen Zeitperiode nur eine begrenzte Strommenge produzieren können, wie Pumpspeicherwerke. An der deutschen Börse wird dieses Gebotsformat jedoch noch nicht genutzt (siehe Grimm, Ockenfels und Zoettl, im Erscheinen, zum Design der europäischen Strombörse EEX, Ockenfels 2007 zu dem Einfluss strategischen Verhaltens auf Strommarktdesign und Stoff 2002 für eine allgemeine Einführung in die Designproblematik bei Strommärkten).

3. Komplexitäten im Verhalten

Bisher war vornehmlich von institutionellen Details die Rede. Angewandtes Marktdesign muss sich jedoch gleichermaßen mit verhaltenswissenschaftlichen Eigenheiten auseinandersetzen – jedenfalls soweit sie sich als relevant herausstellen.

Institutionen spielen für ökonomisches Verhalten eine zentrale Rolle. Der Grund ist, dass Institutionen Anreize beeinflussen, und dass Entscheidungsträger auf Anreize reagieren. Sie tun dies jedoch nicht immer auf rationale (oder egoistische) Weise – doch auch nicht irrational oder chaotisch. Vielmehr liegt dem menschlichen Verhalten eine dem Menschen eigene Rationalität zugrunde, die es oft ermöglicht, Verhalten systematisch zu prognostizieren. Andernfalls wäre Marktdesign als angewandte Disziplin nicht denkbar.

Für Marktdesign relevante Abweichungen vom Standardmodell des homo oeconomicus treten in mindestens zweifacher Weise auf. Erstens spielt die *Motivation* der Agenten, einschließlich

⁵ Negative Preise wären nicht nötig, wenn es über die Stromstundenkontrakte hinaus unbeschränkte Flexibilität bei konditionalen Geboten, erstbeste Optimierungsalgorithmen bei der Preisfindung und keine Unsicherheit gäbe, denn unter akkurater Einbeziehung aller komplementären Kostenbestandteile wäre kein Anbieter bereit, insgesamt negative Preise für ein Stromerzeugungspaket zu akzeptieren.

sozialer Präferenzen, zuweilen eine wichtige Rolle. Beispiele sind das Design innerbetrieblicher Anreizsysteme und von Reputationssystemen; davon wird im Folgenden noch die Rede sein. Zweitens gibt es auch Beschränkungen bei der menschlichen *Kognition*. Naive inkrementelle Verhaltensmuster bei eBay, die Unfähigkeit, auch mit schnellen Computern kombinatorische Probleme in Strommärkten in kurzer Zeit lösen zu können, die nur eingeschränkt rationale Informationsverarbeitung in Märkten (die aktuelle Finanzmarktkrise ist womöglich ein gutes Beispiel; siehe Ockenfels und Selten 2005 für ein anderes, kontrolliertes Beispiel), sowie die Tatsache, dass weder Wissenschaftler noch Studierende immer alle Herausforderungen und Aufgaben schnell und richtig bewältigen können, sind Beispiele für die nur eingeschränkten Rechenkapazitäten realer Entscheidungsträger (siehe Kagel und Roth 1995 sowie Camerer 2003 für Überblicke über die Literatur eingeschränkt rationalen Verhaltens).

Zuweilen ist die Meinung zu hören, dass sich soziale Präferenzen und eingeschränkte Rationalität zumindest in realen Märkten mit Händlern, die sich selbst in den Markt selektiert haben und dort Erfahrungen sammeln konnten, nicht weiter auswirken und daher irrelevant sind. Diese Vorstellung erscheint zwar naheliegend, ist aber oft falsch. Die Theorie selbst legt nahe, dass bereits geringfügige Abweichungen Einzelner vom Rationalverhalten oder von der Eigennutzhypothese das Verhalten *aller* und somit auch das Marktergebnis auf den Kopf stellen können – zumindest in Situationen, bei denen sich individuelle Entscheidungen gegenseitig verstärken (also unter „strategischer Komplementarität“) – und die experimentelle Evidenz bestätigt dies nachdrücklich. (Beispiele findet man etwa bei Kreps und Wilson 1982, Bolton und Ockenfels 1998, 2000, 2008b, Fehr und Schmidt 1999, Fehr und Tyran 2005 und 2008, Nagel 1995, sowie in Kapitel 5). Auch ist die Vermutung, dass sich Experten stets rationaler verhalten als zum Beispiel Studierende, oder dass hinreichend hohe Anreize Irrationalitäten verhindern, ist oft widerlegt worden (siehe Bolton, Ockenfels und Thonemann 2008 und Fehr, Fischbacher und Tougareva 2002 sowie die dort zitierte Literatur).

Marktarchitekten müssen sich daher auch mit der Robustheit von Mechanismen bezüglich Abweichungen vom Rationalverhalten beschäftigen. Auch hier sind die Details wichtig. Ein C2C-Internetmarkt wie eBay kann beispielsweise nicht dieselben Ansprüche an die kognitiven Fähigkeiten seiner Teilnehmer unterstellen, wie die Bundesnetzagentur bei einer Versteigerung von Frequenzen. Im ersteren Fall müssen die Regeln so ausgestaltet sein, dass naive Strategien von erfahrenen Marktteilnehmern nicht ausgenutzt werden können; Robustheit der Marktergebnisse gegenüber nicht-rationalem Verhalten und bei Interaktion von strategischem und naivem Verhalten wird zu einer Kernherausforderung. Im zweiten Fall kann man oft davon ausgehen, dass Strategieexperten zu Rate gezogen werden, die Fehler soweit möglich vermeiden helfen.⁶

Obwohl es in den vergangenen Jahren große Fortschritte bei der Erforschung menschlichen ökonomischen Verhaltens gegeben hat, steht die Umsetzung der Ergebnisse in Anreizsysteme und Märkte erst am Anfang. Erste Arbeiten im Bereich „behavioral mechanism design“

⁶ Eine zu Marktdesign komplementäre Disziplin ist „cognitive engineering“, in der es – vereinfacht gesagt – darum geht, es Menschen leicht zu machen, sich optimal zu verhalten (siehe Roth, Patterson und Mumaw 2002).

finden sich in Kapitel 5, bei Choi, Laibson und Madrian (im Erscheinen), Sunstein und Thaler (2008) sowie Ockenfels, Sliwka und Werner (2008). Letztere zeigen in empirischen Untersuchungen, dass soziale Vergleiche und andere Referenzpunkte bei der Bewertung und Wahrnehmung von Bonus-Auszahlungen für Manager in einem DAX-Unternehmen eine hochsignifikante Rolle spielen, während die absolute Höhe der Auszahlung keine erkennbare Rolle spielt. Dies steht im Einklang mit verhaltenswissenschaftlicher Forschung (Kahneman und Tversky 1979, Bolton und Ockenfels 2000, 2005, Fehr und Schmidt 2000, 2006, Frank 1984, Frank und Sunstein 2001). Die Autoren zeigen dann, wie die Wirkung der Bonus-Zahlungen durch die Transparenz und Präsentation des Bonussystems systematisch beeinflusst werden kann.

4. Die Rolle der Experimentellen Wirtschaftsforschung

Die Experimentelle Wirtschaftsforschung kann als komplementäre Methode zur Theorie und Feldforschung beim angewandten Marktdesign nützlich sein. Insbesondere können Experimente in dreifacher Hinsicht eine Brückenfunktion zwischen Theorie und Praxis bereitstellen:

Theoriebildung und Faktensammlung. Marktdesign erfordert ein Verständnis der kausalen Zusammenhänge von Institutionen, Verhalten und Marktergebnissen. Theorie offenbart oder postuliert Kausalitäten. Eine Aufgabe der Experimentellen Wirtschaftsforschung ist, die theoretisch vermuteten Zusammenhänge und zugrunde liegenden Annahmen auf ihren empirischen Gehalt zu testen (Kagel und Roth 1995). Ein Mechanismus, der sich schon im Labor – also unter exakt den theoretisch postulierten Bedingungen – nicht funktioniert, sollte für die Praxis zumindest mit Vorsicht verwendet werden.

Doch Experimente können auch die Theorie beeinflussen. Tatsächlich ist eine zweite wichtige Aufgabe der Experimentellen Wirtschaftsforschung, Fakten zu sammeln und Verhaltensphänomene und -muster zu identifizieren. Diese können dann Input für deskriptiv relevante Theoriebildung liefern, insbesondere für die Verhaltensökonomik (Schmidt, dieses Heft).

Windkanal. Während Experimente zum Testen und zur Entwicklung von Theorie sicherstellen, dass die zugrunde liegenden Annahmen exakt umgesetzt werden, ist es vornehmliche Aufgabe der Windkanaltests, die Praxistauglichkeit von Mechanismen zu erproben. Viele erdachte Marktregeln, Prozeduren oder Algorithmen gibt es (noch) nicht im Feld, so dass keine empirischen Erfahrungen vorliegen, die die prinzipielle Funktionsfähigkeit belegen könnten. Es liegt daher nahe, einen Prototyp des geplanten Mechanismus im Labor zu testen. Die Erarbeitung des Prototyps und die Erprobung im Windkanal sind nützlich sowohl bezüglich der technischen als auch der ökonomischen Robustheit. Erstens können dabei Probleme der praktischen Implementierbarkeit der Marktregeln identifiziert und die Systemzuverlässigkeit geprüft werden. Und zweitens erlaubt ein Windkanaltest, das neue System auch unter Stressbedingungen zu untersuchen – also zum Beispiel bei ungünstigen Präferenzkonstellationen, Kostenstrukturen, Informationsszenarien oder Teilnehmerzahlen.

Auch können Aspekte, die in der theoretischen Betrachtung nicht adäquat eingefangen werden, frühzeitig identifiziert und so Probleme vermieden werden. Dazu gehören beispielsweise der Einfluss der Kommunikationskanäle und der endogenen Dynamik auf Verhalten und Marktergebnisse. Es gibt viele Beispiele für experimentelle Windkanalversuche; siehe etwa Chen und Ledyard (2006) für einen Überblick – Kapitel 5 beschreibt ein weiteres. Angemerkt sei, dass Feldexperimente dabei eine zunehmend wichtige Rolle einnehmen. Feldexperimente können die Wirkungsweise spezifischer Regelwerke einerseits in kontrollierter Form und andererseits zugleich in praxisnahen Szenarien in natürlicher Umgebung untersuchen. Sie können dadurch die Lücke zwischen Labor und Feld schließen, indem sie das Labor ins Feld beziehungsweise das Feld ins Labor bringen (siehe List 2006, Gneezy und List 2006, sowie Bolton und Ockenfels 2008 für Beispiele).

Kommunikation. Eine letzte Brückenfunktion der Experimentellen Wirtschaftsforschung ist die zwischen Ökonomen und Praktikern. Marktdesign ist eine inhärent interdisziplinäre Aufgabe, doch typischerweise sind nur wenige Manager, Politiker, Marketingexperten, Juristen oder Programmierer dazu in der Lage, den wirtschaftstheoretischen Ansätzen, Überlegungen und Beweisen in gebotener Tiefe folgen zu können. Es zeigt sich, dass die Laborforschung hervorragend geeignet ist, ökonomische Zusammenhänge ‚am eigenen Leib‘ erfahrbar und begreifbar zu machen (Asker et al. 2004). Zudem gilt, dass andere Disziplinen, die am Marktdesign zuweilen beteiligt sind, zwar nicht mit der Wirtschaftstheorie, aber mit experimentellen Methoden vertraut sind.

5. Vom Labor ins Feld: Design eines neuen Reputationssystems für eBay

In diesem Kapitel möchte ich den Nutzen der Experimentellen Wirtschaftsforschung am Beispiel des Designs von eBays neuem Reputationssystem 2007 illustrieren; alle nachfolgenden Daten und Ergebnisse stammen von Bolton, Greiner und Ockenfels (2008; siehe auch Greiner und Ockenfels 2008), wo weitere Hintergründe und Details zu finden sind.

Die Ausgangslage: geringer Informationswert von Reputationsinformation. Transaktionen auf eBay finden typischerweise zwischen Marktteilnehmern statt, die keine gemeinsame Vergangenheit und keine gemeinsame Zukunft haben (Resnick und Zeckhauser 2002), die anonym und geographisch voneinander entfernt kommunizieren und interagieren, und die nur eingeschränkt auf die rechtliche Durchsetzbarkeit ihrer Ansprüche hoffen können. Um die dadurch induzierten Probleme moralischen Risikos und adverser Selektion zu bekämpfen, gibt es bei eBay (und auf vielen anderen Internet-Marktplattformen) einen Reputationsmechanismus, der es nach jeder Transaktion sowohl Käufern als auch Verkäufern erlaubt, eine Bewertung über den anderen Transaktionspartner abgeben. Bis Frühjahr 2007 bestand diese Bewertung aus einem „positiv“/„neutral“/„negativ“-Urteil sowie einer kurzen verbalen Erläuterung, die aggregiert und zukünftigen Transaktionspartnern zur Verfügung gestellt werden. Vertrauenswürdigen Verkäufern wird so die Möglichkeit gegeben, sich eine gute Reputation aufzubauen, während nicht-vertrauenswürdige Verkäufer aus dem Markt gedrängt werden können.

Tatsächlich scheint das Reputationssystem recht gut zu funktionieren. Im Jahre 2007 wurden auf eBay in 70% der möglichen Fälle Bewertungen (nämlich ca. 4 Millionen pro Tag) abgegeben, die auch ökonomisch nachweisbare Effekte erzeugen: bessere Bewertungen führen zu einer erhöhten Verkaufsquote und höheren Preisen. Empirische Studien sowie kontrollierte Feld- und Laborexperimente kreieren dabei ein robustes und weitgehend konsistentes Bild des ökonomischen Einflusses von eBays Reputationssystem.⁷

Zugleich ist ein prinzipielles Problem von Bewertungssystemen die Tendenz der Teilnehmer, übermäßig positive Bewertungen abzugeben. Dies gilt gleichermaßen für die Bewertung von Mitarbeitern im Rahmen innerbetrieblicher Anreizsysteme, für das Versagen von Wirtschaftsprüfern und Rating-Agenturen als eine der Ursachen der Finanzmarktkrise im Herbst 2008, als auch für eBays traditionelles Bewertungssystem. Tatsächlich sind die Bewertungen auf eBay so stark positiv, dass sie kaum eine Diskriminierung der Verkäufer zulassen. Bolton, Greiner und Ockenfels (2008) finden in ihrem Datensatz mit mehr als einer Millionen Beobachtungen einen Anteil von nur 1,4% negativen Bewertungen. Resnick und Zeckhauser (2002) und Dellarocas und Wood (2008) finden mit ca. 0,6% sogar geringere Anteile an negativen Bewertungen. Diese Zahlen scheinen schon aufgrund der im eBay Feedback Forum berichteten alltäglichen Erfahrungen nicht der tatsächlichen Zufriedenheit der Teilnehmer zu entsprechen. Dellarocas und Wood (2008) schätzen auf der Grundlage eines strukturellen statistischen Modells und eines Felddatensatzes, dass nur 79% der Käufer mit dem Ergebnis ihrer Transaktion zufrieden sind. Bolton und Ockenfels (2008b) können in einem kontrollierten Feldexperiment mit realen eBay-Händlern auf der realen Plattform nachweisen, dass unzufriedene Käufer lieber kein Feedback abgeben als ein negatives.

Das Problem: Reziprozität bei der gegenseitigen Bewertung. Ein Grund für den hohen Anteil an positiven Bewertungen – und damit für den verminderten Informationswert des Bewertungssystems – ist (strategische) Reziprozität in der Abgabe der Bewertungen: Bewertungen werden nicht unabhängig voneinander gegeben, sondern werden (strategisch) eingesetzt, um entweder eine positive Bewertung des Transaktionspartners über einen selbst

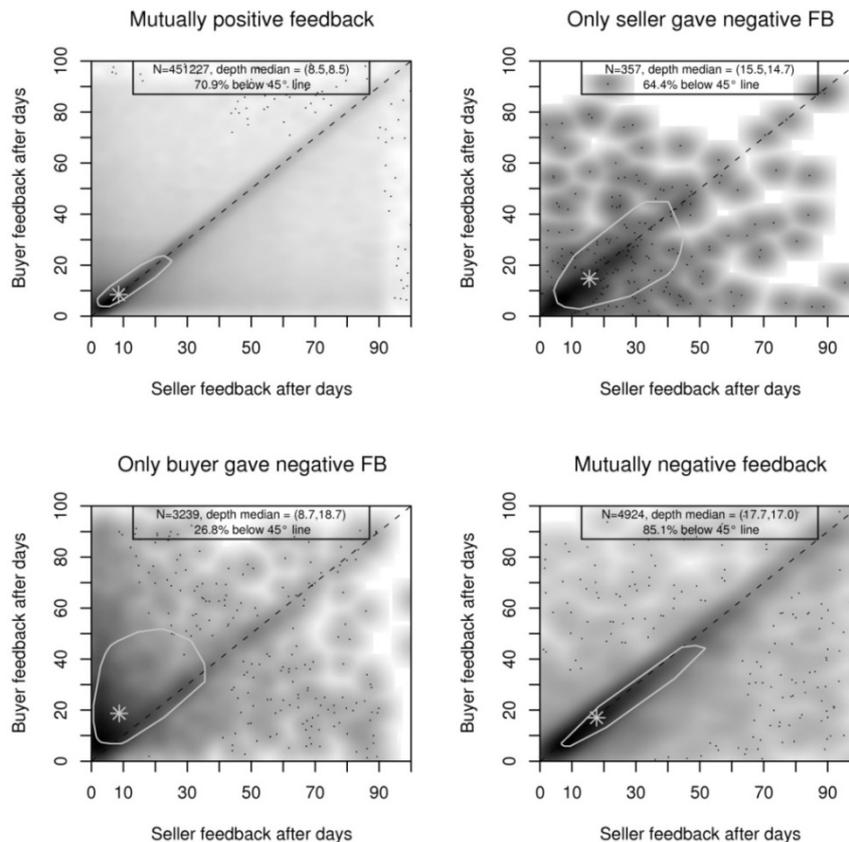
⁷ So finden Resnick und Zeckhauser (2002) und Ockenfels (2003) in Felddaten einen signifikanten Zusammenhang zwischen Bewertungsprofil und Erfolgsquote eines Verkäufers. Melnik und Alm (2002), Livingstone (2005), Houser und Wooders (2006) und Lucking-Reiley et al. (2007) untersuchen Münz- bzw. Elektronikauktion auf eBay und finden signifikante Preisaufschläge für positive Bewertungsprofile. Livingstone (2005) zeigt dass jedoch die marginale Rendite aus positiven Bewertungen mit steigender Zahl der Bewertungen fällt. Im Rahmen eines kontrollierten Feldexperiments arbeiten Resnick et al. (2006) mit einem erfahrenen eBay-Verkäufer zusammen, der unter seinem eigenen und mehreren neu angelegten Profilen Paare von ähnlichen Produkten in qualitativ gleichwertiger Darstellung anbot. Die Ergebnisse bestätigen den positiven Zusammenhang zwischen positiven Bewertungen und Preisaufschlägen. Bolton und Ockenfels (2008b) präsentieren eine kontrollierte Feldstudie, die auch die Produktionsstufe von Bewertungen mit einschließt. Experimentelle Studien, die derlei Ergebnisse unter hochkontrollierten Laborbedingungen replizieren und zusätzlich zum Beispiel den Einfluss von Interaktionsmustern, Informationsflüssen, Wettbewerb der Verkäufer und verschiedene Designaspekte auf Reputationsaufbau und Vertrauen isolieren findet man bei Keser (2003), Bolton, Katok und Ockenfels (2004a,b), Bolton, Loebbecke und Ockenfels (2008), sowie Bolton und Ockenfels (2008a, forthcoming), in der soziologischen Literatur etwa bei Buskens und Raub (forthcoming) und Raub und Buskens (2004). Güth und Ockenfels (2000, 2005) sowie Güth, Mengel und Ockenfels (2007) simulieren eingeschränkt rationalen Reputationsaufbau in anonymen Märkten mit Hilfe von evolutionstheoretischen Methoden.

auszulösen, oder um eine negative Bewertung zu verhindern.⁸ Bolton, Greiner und Ockenfels (2008) beobachten beispielsweise, dass die Wahrscheinlichkeit gegenseitiger Bewertungen ökonomisch und statistisch signifikant höher ist, als man unter der Nullhypothese erwarten sollte, dass die Entscheidung zur Abgabe einer Bewertung unabhängig getroffen wird. Auch gilt, dass nach einer positiven Bewertung des Käufers der Verkäufer zu 99.9% ebenfalls eine positive Bewertung zurückgibt, und nach einer negativen Bewertung die Antwort zu 90% ebenfalls negativ ist. Die wohl überzeugendste Evidenz für Reziprozität leitet sich jedoch aus der endogenen Dynamik der Zeitpunkte der Bewertungen ab.

Wie Abbildung 1 illustriert gibt in den meisten Fällen der Käufer zuerst seine Bewertung ab, und der Verkäufer reagiert darauf. In den allermeisten Fällen sind beide Bewertungen positiv (Grafik oben links). Tatsächlich schafft die reziproke Beziehung zwischen den Transaktionspartnern Anreize, sich gegenseitig positiv zu beurteilen: „wenn Du nett zu mir bist, bin ich auch nett zu Dir“. Doch wo positive Reziprozität ist, ist meistens auch negative Reziprozität: In den (weit weniger) Fällen, in denen sich beide Transaktionspartner negativ beurteilen (Grafik unten rechts), scheint es, als ob der Verkäufer nur auf die Bewertung des Käufers gewartet hat, und sich dann sofort mit einer negativen Bewertung revanchiert; die meisten Beobachtungen liegen knapp unterhalb der Diagonalen. Während positive Reziprozität Anreize schafft, sich überhaupt gegenseitig zu beurteilen, und somit eine gute Reputation aufzubauen, schreckt negative Reziprozität zukünftige Negativbewertungen ab und kann somit dem Schutz der Reputation dienen. Indirekte Evidenz dafür ist in der Grafik unten links zu sehen: Wenn letztendlich nur der Käufer negativ, der Verkäufer jedoch positives Feedback gegeben hat (was selten genug passiert), wird beobachtet, dass die meisten Käufer solange warten, bis der Verkäufer seine Bewertung zuerst losschickt. Dies ist auch klug, muss er doch andernfalls mit einer prompten „Rachebewertung“ rechnen.

⁸ Siehe auch Bajari und Hortaçsu (2004), Cabral und Hortaçsu (2006), Chwelos und Dhar (2007), Dellarocas und Wood (2008), Klein, Lambert, Spagnolo und Stahl (2007), Resnick, Zeckhauser, Friedman, und Kuwabara (2000), und Resnick und Zeckhauser (2002), neben anderen, die ebenfalls die strategische Abgabe von Bewertungen auf eBay diskutieren.

Abbildung 1: Die zeitliche Abfolge von Bewertungen auf eBay



Anmerkung: Datenquelle: Daten von Bolton, Greiner und Ockenfels (2008). Die Grafik enthält ca. 460.000 Beobachtungen bei denen sowohl Käufer als auch Verkäufer eine Bewertung abgegeben haben. „Negative“ enthält sowohl negative, neutrale als auch letztendlich zurückgenommene Bewertungen. Die Graustufen stellen die relative Dichte an Beobachtungen innerhalb des 0.1-Tages-Quadrates dar. Die Taschen enthalten die 50% Beobachtungen mit der höchsten Dichte, der Stern zeigt die Stelle mit der höchsten Konzentration. In jeder Untergrafik sind die 200 am weitesten vom Zentrum entfernt liegenden Beobachtungen als Punkte gekennzeichnet.

Die Daten zeigen, dass Bewertungen nicht unabhängig, sondern strategisch reziprok vergeben werden. Im Ergebnis führt positive Reziprozität in Verbindung mit abschreckenden Rachebewertungen zu übermäßig positiven Bewertungen und erschwert somit den Aufbau von Vertrauen und Vertrauenswürdigkeit.

Design. Designoptionen zur Bekämpfung des Problems beeinflussen die *Dynamik* der Bewertungen und die *Kommunikationskanäle*. Zwei solcher Systeme sollen basierend auf Bolton, Greiner und Ockenfels (2008) hier vorgestellt werden. In einem *blinden* Bewertungssystem werden beide Händler erst dann über die Bewertung des anderen Transaktionspartners informiert, wenn beide Bewertungen abgegeben wurden oder der Bewertungszeitraum verstrichen ist. Die Konsequenz ist, dass die Händler ihre Bewertung nicht auf die Bewertung des Transaktionspartners konditionieren können, und sequenzielle Reziprozität und strategisches Timing damit ausgeschlossen sind. Die Erwartung ist, dass ein solches System zu zuverlässigeren Bewertungen führt. Ein blindes System hat jedoch auch Nachteile. Zum Beispiel ist durch die Vermeidung negativer Reziprozität

(Rachebewertungen) auch positive Reziprozität nicht mehr möglich. Letztere trägt jedoch signifikant dazu bei, dass überhaupt Bewertungen abgegeben werden und die Marktteilnehmer eine Reputation aufbauen können.

In einem *hybriden* System werden zusätzlich zu dem offenen klassischen System detailliertere Bewertungen eingeführt, die nur vom Käufer über den Verkäufer abgegeben werden können, und nicht auf einen spezifischen Käufer zurückzuführen sind („detailed seller ratings“, DSR). Rachebewertungen werden auf eBay vor allem von Verkäufern als Drohmittel, Bestrafung und Disziplinierungsmaßnahme verwendet. Das hybride DSR-System erlaubt den Käufern nun, neben der traditionellen Bewertung auch eine zusätzliche Bewertung abzugeben, die nicht reziprok beantwortet werden kann. Dadurch bleibt die Möglichkeit der positiven Reziprozität im offenen Teil des hybriden Systems (dem klassischen System) erhalten, während Käufer die Möglichkeit haben, Kritik in differenzierterer Form als bisher in zusätzlichen, einseitigen und verdeckten Bewertungen zu äußern.

Experimente. Experimente spielten bei der Neugestaltung von eBays Reputationssystem 2007 in mehrfacher Hinsicht eine entscheidende Rolle (für Details siehe Bolton, Greiner und Ockenfels 2008).⁹

Experimente im Labor und Feld haben komplementär zu empirischen Analysen von Felddaten und Umfragen unter eBay-Händlern die Funktionsweise, -fähigkeit und -defizite des bestehenden Systems in kontrollierter Umgebung diagnostiziert und quantifiziert (siehe auch Fußnote 7).

Natürliche Experimente, bei denen weltweit operierende reale Online-Marktplattformen untersucht wurden, die sich in ihrem Reputationssystem unterscheiden, legten nahe, dass das Systemdesign signifikant auf das Bewertungsverhalten einwirkt. Zum Beispiel gibt es auf eBays Plattform in Brasilien, die ein blindes Reputationssystem einsetzt, signifikant mehr negative Bewertungen als in den USA und in Europa. Und Rent-a-coder.com, eine Marktplattform, die von einem eBay-ähnlichen auf ein blindes System umgestiegen ist, verzeichnete nach dem Systemwechsel eine signifikant geringere Korrelation der Bewertungen. Beides legt nahe, dass ein blindes Bewertungssystem die strategisch reziproken Abhängigkeiten verringert und somit den Informationswert der Reputation vergrößert.

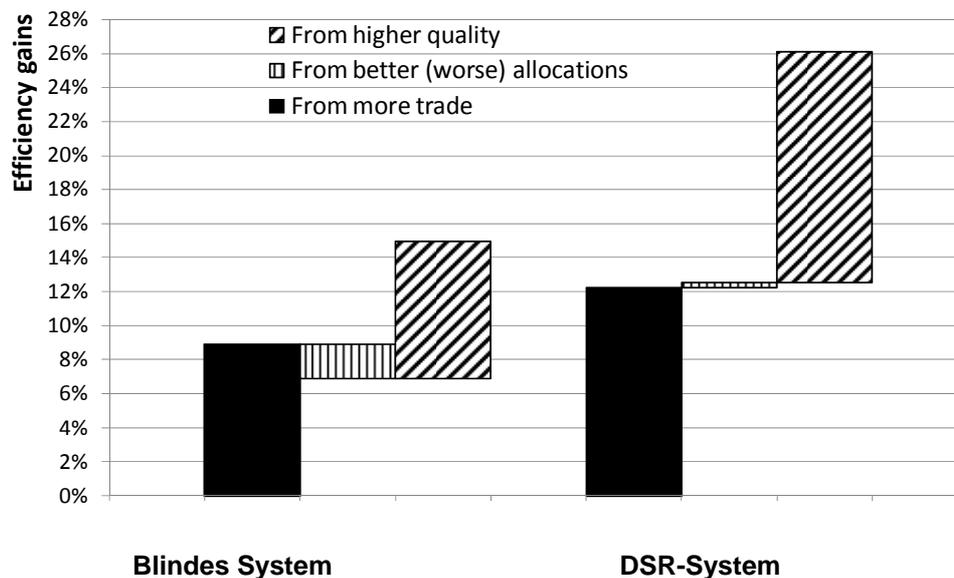
Felddaten lassen jedoch in der Regel keine eindeutigen Rückschlüsse auf kausale Zusammenhänge zu. Zum Beispiel könnte die in Brasilien beobachtete erhöhte Neigung zu negativen Bewertungen mit einem schlechter funktionierendem Kreditkarten- oder Postsystem, oder auch durch eine größere, kulturell bedingte Toleranz gegenüber negativen Bewertungen zu erklären sein. Kontrollierte Laborexperimente (Windkanaltests) können

⁹ In den Laborexperimenten handeln Teilnehmer auf einem an eBay angelehnten Auktionsmarkt mit unabhängigen privaten Bewertungen über jeweils 60 Runden. Allen experimentellen Anordnungen liegt eine Zweitpreisauktion zugrunde, wie sie auch auf eBay stattfindet (Ockenfels und Roth 2006). Nach der Auktion kann der Verkäufer über die Qualität des Produktes entscheiden, als ganze Zahl zwischen 1 und 100. Jeder Qualitätspunkt kostet dem Verkäufer eine Geldeinheit und erbringt dem Käufer 1% seiner Wertschätzung. Dadurch wird das moralische Risiko auf eBay induziert. Nach abgeschlossenem Handel startet die Bewertungsphase, die sich je nach experimenteller Anordnung unterscheidet, und dem System von eBay ähnelt.

jedoch komplementär nachweisen, dass die in der unkontrollierten Feldumgebung beobachteten Effekte kausal auf das System zurückgeführt werden können. (Die Robustheit der Effekte wird durch die Tatsache unterstrichen, dass nicht nur qualitative sondern auch viele quantitative Bewertungsmuster auf eBay im Labor repliziert werden konnten.) Zudem kann nachgewiesen werden, dass die positiven Effekte insbesondere dadurch erzeugt werden, dass die in den neuen Bewertungssystemen enthaltenen Informationen tatsächlich eine bessere Prognose des zu erwartenden Verkäuferverhaltens ermöglichen.

Letztlich liegt das eigentliche Designziel in der ökonomischen Effizienz des *Handels* – und nicht in der Funktionsfähigkeit des Reputationssystems. Diese kann jedoch mangels Daten über Zahlungsbereitschaften und Kosten nicht (ohne weiteres) mit Felddaten gemessen werden. Laborexperimente können dagegen den Einfluss des Reputationssystems auf die ökonomisch relevanten Daten durch die Kontrolle von Zahlungsbereitschaften und Kosten exakt messen. Dabei zeigt sich, dass beide Systemalternativen zu einer signifikanten Steigerung der Produktqualität und Vertrauenswürdigkeit von Verkäufern, und zugleich entsprechend auch zu einer erhöhten Bietbereitschaft der Käufer führen.

Abbildung 2: Effizienzgewinne in experimentellen Märkten mit blindem und hybridem Bewertungssystem



Anmerkung: Quelle: Bolton, Greiner und Ockenfels (2008). Überlappende Balken kennzeichnen Effizienzverluste. Alle Effizienzgewinne (-verluste) sind relativ zur Anordnung mit klassischem Bewertungssystem.

Abbildung 2 zeigt die Effizienzgewinne innerhalb der verschiedenen experimentellen Märkte und ihre Quellen. Die Effizienz steigt sowohl durch mehr Handel (also mehr Vertrauen) und bessere Allokationen als auch durch höhere Produktqualität (also mehr Vertrauenswürdigkeit), und sowohl Käufer als auch Verkäufer profitieren davon. Damit weist

das Experiment nach, dass sich das verbesserte Reputationssystem auch materiell in erhöhter Markteffizienz messen lässt.¹⁰

Die theoretischen Prognosen zur Performanz der Reputationssysteme sind notorisch schwach, da wir es mit Situationen mit wiederholter Interaktion zu tun haben.¹¹ Zudem abstrahiert die Theorie typischerweise von Pfadabhängigkeiten, die in der Praxis jedoch oft die Optionen für Marktdesign beschneiden; eBay wäre nicht gut beraten, das alte System einem neuen zu opfern, wenn dadurch viele Millionen Nutzer ihre Reputation verlieren würden oder die Vergleichbarkeit alter und neuer Bewertungen gefährdet wäre. Schließlich kann die Berücksichtigung historischer Daten auch helfen, ausgehend vom Status Quo *Verhaltensänderungen* zu prognostizieren (oder Probleme der Gleichgewichtsauswahl zu adressieren). Tatsächlich führen die Experimente zusammen mit den Feld- und Umfragedaten zu vergleichsweise robusten Prognosen.

Kontrollierte Feldexperimente sowie Pilotexperimente in einigen europäischen und außereuropäischen Ländern, in denen ein neues Bewertungssystem durch eBay auf die jeweilige Plattform aufgespielt wurde, schließen letztlich die letzten Lücken zwischen Theorie, Laboratorium und Praxis.

Schließlich hat sich die Entwicklung und Analyse der Laborexperimente als ausgezeichnete Plattform erwiesen, um Managern, Juristen, Programmierern, Marketingexperten, Marktteilnehmern und anderen involvierten Personen komplexe ökonomische Zusammenhänge zu illustrieren,¹² sowie die Vor- und Nachteile der verschiedenen Systeme zu diskutieren und gegeneinander abzuwägen. Es steht zu bezweifeln, dass wirtschaftstheoretische Resultate alleine – wenn sie denn robuste Einsichten liefern können – bei Praktikern dieselbe Wirkung erzielen können.

Ergebnis. Das DSR-System wurde auf der Basis der vorgenannten Ergebnisse von Gedanken-, Labor- und Feldexperimenten im Frühjahr 2007 unter dem Namen „Feedback 2.0“ von eBay weltweit eingeführt. Wie Bolton, Greiner und Ockenfels (2008) zeigen, sind die detaillierten Verkäuferbewertungen auf eBay nun deutlich differenzierter (und negativer) als die konventionellen Bewertungen. Teilweise wird auch beobachtet, dass Käufer negative detaillierte, verdeckte Bewertungen hinter positiven konventionellen, offenen Bewertungen verbergen. Zudem gibt es auch indirekte Evidenz für positive ökonomische Effekte.

¹⁰ Das blinde System führt zu einer deutlich geringeren Anzahl an Bewertungen, insbesondere für Verkäufer. Eine Ursache ist der Wegfall der positiven Reziprozität im blinden System, was letztlich zu dem etwas schlechteren Ergebnis im Vergleich zu DSR beiträgt.

¹¹ Tatsächlich würde die Theorie unter der Annahme, dass alle Marktteilnehmer egoistisch und rational sind und beides gemeinsames Wissen ist, für das Experiment (mit endlich wiederholter Interaktion) keinerlei Handel voraussagen – unabhängig davon, wie das Reputationssystem aussieht. Doch das Resultat ist nicht robust. Schon wenn Rationalität nicht mehr gemeinsames Wissen ist, ist Kooperation im Gleichgewicht erreichbar. Zudem lässt sich praktisch jeder Spielpfad durch minimale Eingriffe in die Nutzenfunktionen einiger Marktteilnehmer im spieltheoretischen Gleichgewicht erzeugen. Unendlich wiederholte Spiele werden auch ohne solche Abweichungen von Standardannahmen durch eine Vielzahl von Gleichgewichtsergebnissen geplagt.

¹² Dazu gehören beispielsweise Endogenitätseffekte: auf der einen Seite steigt die Anzahl negativer Bewertungen in einem blinden System, da Rachefeedback weniger gefürchtet werden muss, auf der anderen Seite fällt sie jedoch aufgrund der verbesserten Anreizwirkung und der damit verbundenen besseren Qualität der Verkäuferaktivitäten.

Insgesamt legen die Daten nahe, dass die Einführung des hybriden Systems die Funktionsweise des Bewertungssystems und damit die Effizienz im Markt verbessert hat – wie es von den Experimenten prognostiziert wurde.¹³

Die Studie demonstriert, wie das Zusammenspiel experimenteller mit anderen Methoden auch in einem theoretisch schwierigen Umfeld die relevanten institutionellen Details und Komplexitäten im Verhalten einfangen und einer systematischen Untersuchung zugänglich machen kann. Insbesondere wird gezeigt, wie Reziprozität – eines der wichtigsten sozialen und ökonomischen Verhaltensphänomene – durch Manipulation der Kommunikations- und Informationskanäle effizienzsteigernd gesteuert werden kann.

6. Schlussbemerkung

Für Marktdesign ist die Experimentelle Wirtschaftsforschung ein Instrument an der Schnittstelle von Theorie und Praxis. Kreativ eingesetzt erlaubt sie, den Übergang von Theorie über das Labor ins Feld unter kontrollierten Bedingungen zu analysieren. Die Experimentelle Wirtschaftsforschung kann Fragen beantworten, die Theorie und Felddaten nicht beantworten können, Hypothesen und Kausalitäten testen, die Theorie und Felddaten nahe legen, Fakten und Phänomene sammeln, die Theorie und Feldforschung befruchten können, und schließlich den Marktteilnehmern, Managern und Politikern komplexe ökonomische Zusammenhänge kommunizieren und Hilfestellung bei Designfragen geben. Im Ergebnis profitieren sowohl die Wissenschaft als auch die Praxis.

¹³ Im Frühjahr 2008 veränderte eBay sein Bewertungssystem erneut: seitdem ist es Verkäufern nicht mehr erlaubt, negative Bewertungen abzugeben. Diese Regeländerung transformiert eBays Bewertungsabgabe grundsätzlich in ein einseitiges Reputationssystem, mit einem Element positiver Reziprozität. Bolton, Greiner, Lamouroux und Ockenfels (in Bearbeitung) diskutieren das neu eingeführte sowie weitere alternative Designs.

Literatur

- Ariely, D., Ockenfels, A. and A. E. Roth (2005). An Experimental Analysis of Ending Rules in Internet Auctions. *Rand Journal of Economics*, 36(4), 891-908.
- Asker, J., Grosskopf, B., McKinney, C. N., Niederle, M., Roth, A. E. and G. Weizsäcker (2004). Teaching Auction Strategy Using Experiments Administered via the Internet. *Journal of Economic Education*, 35, 330-342.
- Bajari, P. and A. Hortaçsu (2004). Economic Insights from Internet Auctions. *Journal of Economic Literature*, 42(2), 457-486.
- Baliga, S. and E. Maskin (2003). Mechanism Design for the Environment. In: K.G. Mäler and J. Vincent (Eds.), *Handbook of Environmental Economics*, 1, Amsterdam: North-Holland, 305-324.
- Bolton, G., Greiner, B., Lamouroux, F. and A. Ockenfels (Work in Progress). Feedback 3.0 - An Investigation into Recent Changes of eBay's Feedback System.
- Bolton, G., Greiner, B. and A. Ockenfels (2008). Engineering Trust - Reciprocity in the Production of Reputation Information. Working paper.
- Bolton, G., Katok, E. and A. Ockenfels (2004a). How Effective are Online Reputation Mechanisms? An Experimental Study. *Management Science*, 50(11), 1587-1602.
- Bolton, G., Katok, E. and A. Ockenfels (2004b). Trust Among Internet Traders: A Behavioral Economics Approach. *Analyse und Kritik*, 26, 185-202.
- Bolton, G., Loebbecke, C. and A. Ockenfels (2008). Does Competition Promote Trust and Trustworthiness in Online Trading? An Experimental Study. *Journal of Management Information Systems*, 25(2), 145-169.
- Bolton, G. and A. Ockenfels (1998). Strategy and Equity: An ERC-Analysis of the Güth-van Damme Game. *Journal of Mathematical Psychology*, 42(2), 215-26.
- Bolton, G. and A. Ockenfels (2000). ERC - A Theory of Equity, Reciprocity and Competition. *American Economic Review*, 90(1), 166-93.
- Bolton, G. and A. Ockenfels (2005). A Stress Test of Fairness Measures in Theories of Social Utility. *Economic Theory*, 25(4), 957-82.
- Bolton, G. und A. Ockenfels (2008b). Does Laboratory Trading Mirror Behavior in Real World Markets? Fair Bargaining and Competitive Bidding on Ebay. University of Cologne, Working paper Series in Economics, No. 36.
- Bolton, G. and A. Ockenfels (2008a). Information Value and Externalities in Reputation Building - An Experimental Study. Working paper, University of Cologne.
- Bolton, G. and A. Ockenfels (forthcoming). The Limits of Trust. In: K. S. Cook, C. Snijders, V. Buskens and C. Cheshire (Eds.), *Trust and Reputation*, New York: Russell Sage.
- Bolton, G., Ockenfels, A. and U. Thonemann (2008). Managers and Students as Newsvendors. How Out-of-Task Experience Matters. University of Cologne, Working paper Series in Economics, No. 39.
- Brosig, J. (2006). Communication Channels and Induced Behavior. *Zeitschrift für Betriebswirtschaft*, Special Issue 5, 99-120.

- Brosig, J., Ockenfels, A. and J. Weimann (2003). The Effect of Communication Media on Cooperation. *German Economic Review* 4(2), 217-241.
- Brunner, C., Goeree, J., Holt, C., and J. Ledyard (2007). An Experimental Test of Combinatorial FCC Spectrum Auctions. Working paper.
- Buskens, V. and W. Raub (forthcoming). Rational Choice Social Research on Social Dilemmas: Embeddedness Effects on Trust. In: R. Wittek, T.A.B. Snijders and V. Nee (Eds.), *Handbook of Rational Choice Social Research*, New York: Russell Sage.
- Cabral, L. und A. Hortaçsu (2006). The Dynamics of Seller Reputation: Evidence from eBay. Working paper, University of Chicago.
- Camerer, C. F. (2003). *Behavioral Game Theory*. Princeton: Princeton University Press.
- Chen, Y. and J. Ledyard (2006). Mechanism Design Experiments. Working paper.
- Chen, Y. and T. Sönmez (2006). School Choice: An Experimental Study. *Journal of Economic Theory*, 127, 202-231.
- Choi, J., Laibson, D. and B. Madrian (forthcoming). Reducing the Complexity Costs of 401(k) Participation Through Quick Enrollment (TM). In: D. A. Wise (Ed.), *Developments in the Economics of Aging*. Chicago: University of Chicago Press.
- Dellarocas, C. and C. A. Wood (2008). The Sound of Silence in Online Feedback: Estimating Trading Risks in the Presence of Reporting Bias. *Management Science*, 54(3), 460-476.
- Fehr, Ernst. Dieses Heft.
- Fehr, E., Fischbacher, U. and E. Tougareva (2002). Do High Stakes and Competition Undermine Fairness? Evidence from Russia. IEW Working paper No. 120, University of Zurich.
- Fehr, E. and K. M. Schmidt (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 114, 817-868.
- Fehr, E. and K. M. Schmidt (2006). The Economics of Fairness, Reciprocity and Altruism? Experimental Evidence and New Theories. In: S.-G. Kolm and J. M. Ythier (Eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity*, Amsterdam: Elsevier, 615-691.
- Fehr, E. and J.-R. Tyran (2005). Individual Irrationality and Aggregate Outcomes. *Journal of Economic Perspectives*, 19(4), 43-66.
- Fehr, E. and J.-R. Tyran (2008). Limited Rationality and Strategic Interaction - The Impact of the Strategic Environment on Nominal Inertia. *Econometrica*, 76(2), 353-94.
- Frank, R. H. (1984). Are Workers Paid their Marginal Products? *American Economic Review*, 74(4), 549-571.
- Frank, R. H. and C. R. Sunstein (2001). Cost Benefit Analysis and Relative Position. University of Chicago, Working paper 102.
- Gneezy, U. and J. A. List (2006). Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets using Field Experiments. *Econometrica*, 74(5), 1365-1384.
- Greiner, B. and A. Ockenfels (2008). Vom Labor ins Feld: Die Ökonomik des Vertrauens. Working paper.

- Grimm, V., Ockenfels, A. and G. Zoetl (forthcoming). Strommarktdesign: Zur Ausgestaltung der Auktionsregeln an der EEX. *Zeitschrift für Energiewirtschaft*.
- Güth, W., Mengel, F. and A. Ockenfels (2007). An Evolutionary Analysis of Buyer Insurance and Seller Reputation in Online Markets. *Theory and Decision*, 63(3), 265-82.
- Güth, W. and A. Ockenfels (2000). Evolutionary Norm Enforcement. *Journal of Institutional and Theoretical Economics*, 156(2), 335-347.
- Güth, W. and A. Ockenfels (2005). The Coevolution of Morality and Legal Institutions: An Indirect Evolutionary Approach. *Journal of Institutional Economics*, 1(2), 155-174.
- Houser, D. and J. Wooders (2006). Reputation in Auctions: Theory, and Evidence from eBay. *Journal of Economics & Management Strategy*, 15(2), 353-369.
- Hurwicz, L. (1972). On Informationally Decentralized Systems. In: R. Radner and C. B. McGuire, *Decision and Organization*, Amsterdam: North-Holland, 297-336.
- Hurwicz, L. (1960). Optimality and Informational Efficiency in Resource Allocation Processes. In: K. Arrow and P. Suppes (Eds.), *Mathematical Methods in the Social Sciences*. Stanford: Stanford University Press, 27-47.
- Jehiel, P. and B. Moldovanu (2006). Allocative and Informational Externalities in Auctions and Related Mechanisms. In: R. Blundell, W. Newey and T. Persson (Eds.), *The Proceedings of the 9th World Congress of the Econometric Society*, Cambridge: Cambridge University Press, 102-135.
- Kagel, J. H. and A. E. Roth (1995). *Handbook of Experimental Economics*. Princeton: Princeton University Press.
- Kagel, J. H. and A. E. Roth (2000). The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Experiment. *Quarterly Journal of Economics*, 115(1), 201-235.
- Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-292.
- Keser, C. (2003). Experimental Games for the Design of Reputation Management Systems. *IBM Systems Journal*, 42(3), 498-506.
- Kittsteiner, T. and A. Ockenfels (2006). Market Design: A Selective Review. *Zeitschrift für Betriebswirtschaft*, 5, 121-143.
- Kittsteiner, T. and A. Ockenfels (2008). On the Design of Simple Multi-unit Online Auctions. In: H. Gimpel, N. R. Jennings, G. E. Kersten, A. Ockenfels and C. Weinhardt (Eds.), *Negotiation, Auctions and Market Engineering*, Berlin, Heidelberg: Springer, 68-71.
- Klein, T. J., Lambertz, C., Spagnolo, G. and K. O. Stahl (2007). Reputation Building in Anonymous Markets: Evidence from eBay. Working paper, University of Mannheim.
- Klemperer, P. (2004). *Auctions: Theory and Practice*. Princeton: Princeton University Press.
- Kreps, D.M., and R. Wilson (1982). "Reputation and Imperfect Information." *Journal of Economic Theory*, 27, 253-279.
- Laibson, David. Dieses Heft.
- Ledyard, J. O. (1995). Public Goods: A Survey of Experimental Research. In: J. H. Kagel and A. E. Roth (Eds.), *Handbook of Experimental Economics*, Princeton: Princeton University Press, 111-194.

- List, J. A. (2006). The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions. *Journal of Political Economy*, 114(1), 1-37.
- Livingston, J. (2005). How Valuable Is a Good Reputation? A Sample Selection Model of Internet Auctions. *Review of Economics and Statistics*, 87(3), 453-465.
- Lucking-Reiley, D., Bryan, D., Prasad, N. and D. Reeves (2007). Pennies from eBay: The Determinants of Price in Online Auctions. *Journal of Industrial Economics*, 55(2), 223-233.
- Melnik, M. I. and J. Alm (2002). Does a Seller's eCommerce Reputation Matter? Evidence from eBay Auctions. *Journal of Industrial Economics*, 50(3), 337-347.
- Milgrom, P. (2004). *Putting Auction Theory to Work*, Cambridge: Cambridge University Press.
- Mongell, S. and A. E. Roth (1991). Sorority Rush as a Two-Sided Matching Mechanism. *American Economic Review*, 81, 441-464.
- Müsgens, F. and A. Ockenfels (2006). Marktdesign in der Energiewirtschaft. In: W. Franz, H. J. Ramser and M. Stadler (Eds.): *Umwelt und Energie, Wirtschaftswissenschaftliches Seminar Ottobeuren, 2006*, Tübingen: Mohr Siebeck, 247-272.
- Nagel, R. (1995). Unraveling in Guessing Games: An Experimental Study. *American Economic Review*, 85(5), 1313-1326.
- Ockenfels, A. (2001). Ökonomisches Marktdesign: Neue Herausforderungen in der neuen Ökonomie. In: O. Bäte, C. Smekal and J. Starbatty (Eds.), *Old and New Economy auf dem Weg in eine innovative Symbiose? Dominanz der Prozesse-Flexibilität der Strukturen-Konstanz der ökonomischen Grundregeln*, J. P. Bachem Verlag: Köln, 162-163.
- Ockenfels, A. (2003). Reputationsmechanismen auf Internet-Marktplattformen: Theorie und Empirie. *Zeitschrift für Betriebswirtschaft*, 73(3), 295-315.
- Ockenfels, A. (2007). Strombörse und Marktmacht. *Energiewirtschaftliche Tagesfragen*, 57(5), 44-58.
- Ockenfels, A., Reiley, D. and A. Sadrieh (2006). Online Auctions. In: T. J. Hendershott (Ed.), *Handbooks in Information Systems I, Handbook on Economics and Information Systems*, 571-628.
- Ockenfels, A. and A. E. Roth (2006). Late and Multiple Bidding in Second-Price Internet Auctions: Theory and Evidence Concerning Different Rules for Ending an Auction. *Games and Economic Behavior*, 55, 297-320.
- Ockenfels, A. and A. E. Roth (2002). The Timing of Bids in Internet Auctions: Market Design, Bidder Behavior, and Artificial Agents. *Artificial Intelligence Magazine*, 23(3), 79-88.
- Ockenfels, A. and R. Selten (2000). An Experiment on the Hypothesis of Involuntary Truth-Signalling in Bargaining. *Games and Economic Behavior*, 33, 90-116.
- Ockenfels, A. and R. Selten (2005). Impulse Balance Equilibrium and Feedback in First Price Auctions. *Games and Economic Behavior*, 51, 155-170.
- Ockenfels, A., Sliwka, D. and P. Werner (2008). Designing Bonus Schemes for Managers. Work in progress.

- Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.
- Raub, W. and V. Buskens (2004). Spieltheoretische Modellierungen und Empirische Anwendungen in der Soziologie. In: A. Diekmann (Ed.), *Methoden der Sozialforschung, Kölner Zeitschrift für Soziologie und Sozialpsychologie, Special Issue 44/2004*, Wiesbaden: Verlag für Sozialwissenschaften.
- Resnick, P. and R. Zeckhauser (2002). Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. *Advances in Applied Microeconomics*, 11, 127-157.
- Resnick, P., Zeckhauser, R., Friedman, E. and K. Kuwabara (2000). Reputation Systems. *Communications of the ACM*, 43(12), 45-48.
- Resnick, P., Zeckhauser, R., Swanson, J. and K. Lockwood (2006). The Value of Reputation on eBay: A Controlled Experiment. *Experimental Economics*, 9(2), 79-101.
- Roth, A. E. (1991). A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K. *American Economic Review*, 81(3), 415-440.
- Roth, A. E. (1990). New Physicians: A Natural Experiment in Market Organization. *Science*, 250, 1524-1528.
- Roth, A. E. (2002). The Economist as Engineer: Game Theory, Experimental Economics and Computation as Tools of Design Economics. *Econometrica*, 70(4), 1341-1378.
- Roth, A. E. (2008). What have we learned from Market Design?. *Economic Journal*, 118(527), 285-310.
- Roth, A. E. and A. Ockenfels (2002). Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet. *American Economic Review*, 92(4), 1093-1103.
- Roth, A. E. and E. Peranson (1999). The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design. *American Economic Review*, 89(4), 748-779.
- Roth, A. E. and X. Xing (1994). Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions. *American Economic Review*, 84, 992-1044.
- Roth, A. E., Sönmez, T., and M. U. Ünver (2004). Kidney Exchange. *Quarterly Journal of Economics*, 119(2), 457-488.
- Roth, E. M., Patterson, E. S. and R. J. Mumaw (2002). *Cognitive Engineering: Issues in User-Centered System Design*. In: J. J. Marciniak (Ed.), *Encyclopedia of Software Engineering (2nd Edition)*, New York: John Wiley and Sons, 163-179.
- Sally, D. (1995). Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992. *Rationality and Society*, 7, 58-92.
- Schmidt, Klaus. Dieses Heft.
- Serrano, R. (2004). The Theory of Implementation of Social Choice Rules. *SIAM Review*, 46, 377-414.
- Stoft, S. (2002). *Power System Economics: Designing Markets for Electricity*. New Jersey: Wiley-IEEE Press.

- Sunstein, R. and R. Thaler (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. London, New Haven: Yale University Press.
- Whinston, Michael D. (2008). *Lectures on Antitrust Economics*, MIT Press.
- Wilson, R. (1985). Reputations in Games and Markets. In: A. E. Roth (Ed.), *Game-Theoretic Models of Bargaining*, Cambridge: Cambridge University Press, 27-62.
- Wilson, R. (2002). Architecture of Power Markets. *Econometrica*, 70(4), 1299-1340.