

## **Vertrauen: Substitut oder Komplement zu formellen Institutionen?**

Von Friedrich L. Sell und Marcus Wiens\*

### **Abstract**

In this paper, we take trust as a specific, cooperative investment. The trustor is mainly interested in the possibility to rely on the trustee („reliance“), while the trustee's object is to receive a trustworthy signal („trust responsiveness“). The value of a trust relationship amounts and equals to not less than the social surplus which it can generate. On the one hand, a social optimal solution is feasible in the case where both parties agree on „liquidated damages“ (perfect jurisdiction). Yet, it is also possible, on the other hand, to achieve a second best solution in a non-perfect jurisdiction world when „trust responsiveness“ comes into play.

### **Zusammenfassung**

In diesem Beitrag fassen wir Vertrauen als spezifische, kooperative Investition auf. Dem Vertrauensgeber geht es dabei darum, sich auf den Vertrauensnehmer verlassen zu können („reliance“), dem Vertrauensnehmer darum, ein vertrauenswürdigen Signal zu empfangen („trust responsiveness“). Der Wert einer Vertrauensbeziehung besteht in nicht weniger als in dem sozialen Überschuss, den sie entstehen lassen kann. Einerseits kann die Zahlung des individuell geregelten Schadensersatzes (perfekte Jurisdiktion) dafür Sorge tragen, die Parteien zur Realisierung der sozial optimalen Lösung zu bewegen. Andererseits ist es auch möglich, dass im Zusammenspiel von „trust responsiveness“ mit imperfekter Jurisdiktion immerhin eine zweitbeste Lösung zustande kommt.

*JEL Classification: K12, D69, D23*

Received: March 30, 2004

Accepted: January 31, 2006

---

\* Wir danken zwei anonymen Gutachtern für wertvolle Hinweise und Verbesserungsvorschläge.

„The essence of market economy is that it involves countless economic exchange transactions. Mutual confidence is what many of those transactions are about. An individual lends money to another individual, relying on that person to return it. This confidence is based on the legal system. If a person fails to meet contractual obligations, there are means to enforce them.“<sup>1</sup>

## 1. Einleitung

Im obigen Zitat zieht Stiglitz Schlussfolgerungen aus seinen Beobachtungen und Erfahrungen während des ökonomischen Transformationsprozesses und weist auf einen wichtigen Teilaspekt des Vertrauens hin: Empirische Studien zeigen nämlich, dass „trust increases with formal institutions“ (Zak / Knack, 2001, 311). Dies lässt sich damit begründen, dass formelle Institutionen häufig gerade dafür geschaffen wurden, um Vertrauen aufzubauen (vgl. ebenda, 316). Warum? Formelle Institutionen sind dazu ermächtigt, „Verträge durchzusetzen“ (vgl. ebenda, 316). Folglich erhöht schon die Existenz formeller Institutionen die Wahrscheinlichkeit (und somit das Vertrauen zur bzw.) für die Einhaltung rechtlicher Normen. Allerdings entspricht dies nicht ganz einem anderen, sehr verbreiteten Verständnis von Vertrauen, wonach „*people may trust each other only for fear of legal punishment. If this were the case, then the effect of trust would simply be a reinterpretation of the role played by legal enforcement*“ (Guiso et al., 2000, 6).“

Guiso et al. führen daher aus, dass Menschen gar keine besondere Notwendigkeit (bzw. Möglichkeit, die Verfasser) hätten, anderen Menschen zu vertrauen, wenn *kein Rechtssystem* und *keine formellen Mechanismen zur Durchsetzung von Verpflichtungen* bestehen würden (Guiso et al., 2000, 6).

Es stellt sich mithin die Frage, wie sich dieser (möglicherweise nur scheinbare) Widerspruch auflösen lässt. Einen wichtigen Hinweis geben Richter / Furubotn: „Das Vorhandensein von Vertrauen, das durch *institutionelle* Garantien geschützt wird, kann zur *Konvention* des Privateigentums und dessen gesellschaftlichen Schutz führen (Richter / Furubotn, 1999, 87). Demnach wäre das Recht kein Substitut, sondern ein (wenn auch eher unvollständiges) Komplement zum Vertrauen. Konventionen können wiederum als Willenserklärungen aufgefasst werden, die Vertrauen von denen verlangen, welche die Adressaten der Willenserklärung sind und wiederum den verpflichten, der sie abgibt (ebenda, 156).

Sowohl die ökonomische Theorie des Rechts als auch die formale Vertragstheorie als Teil der Neuen Institutionenökonomik bietet einen guten Referenzrahmen, um zu untersuchen, ob und – falls ja – wann man das Entstehen bzw.

<sup>1</sup> Stiglitz, 2002, 191.

die Existenz von Vertrauensbeziehungen nur mit bzw. auch ohne die Existenz formeller, zu Sanktionen berechtigter Institutionen erklären kann. Besonders anschauliche Beispiele für solche Institutionen sind Gerichte und deren Urteile. Im Rahmen des Zivilrechts werden Gerichte bei unzureichender Erfüllung oder dem Bruch von Rechtsgeschäften häufig mit Schadensersatzforderungen einer der Parteien beschäftigt. Dabei spielen sogenannte „Schadensersatzregeln“ eine zentrale Rolle. Wir wollen nun in diesem Beitrag zeigen, dass die in diesem Zusammenhang entwickelte ökonomische Theorie des Rechts – im Lichte des oben beschriebenen Widerspruchs – zu wichtigen Rückschlüssen im Hinblick auf das Vertrauensphänomen beitragen kann.

## 2. Der relevante Vertrauensbegriff

Die mittlerweile sehr umfangreich gewordene ökonomische Literatur zum Phänomen des Vertrauens ist weit davon entfernt, über einen einheitlichen Vertrauensbegriff zu verfügen.<sup>2</sup> Soweit es das interpersonale Vertrauen betrifft und nicht das Vertrauen gegenüber Institutionen, wird dieses häufig als bewusste Inkaufnahme des Risikos eines Nichtvollzugs der vom Gegenüber erwarteten Leistung gesehen – durchaus in dem Sinne, sich in gewisser Weise in die Hand des Partners zu begeben:

„Trust in things or people entails the willingness to submit to the risk that they may fail us, with the expectation that they will not, or the neglect or lack of awareness of the possibility that they might“ (Nooteboom, 2002, 45).

Nooteboom spricht in diesem Fall von „trust on a non-calculative basis“ (ebenda). Davon zu unterscheiden wäre „trust on a calculative basis“ (ebenda, 42), das einem Gegenüber aus der Erwartungshaltung heraus gewährt wird, dass der Nutzen daraus, ein solches Risiko einzugehen, sich durchaus „rechnen“ kann. Dies ist dann der Fall, wenn den Kooperationspartner etwa die Furcht vor einem Reputationsverlust bzw. die Aussicht auf Kooperationsgewinne daran hindern, den kurzfristigen Ertrag eines opportunistischen Missbrauchs von Vertrauen zu erbeuten. Die Grundlage für die riskante Vertrauensentscheidung besteht in einer beziehungsinhärenten Anreizkonstellation, die den Partner vom Vertrauensmissbrauch abhält. In der Literatur spricht man in diesem Zusammenhang auch von den „Schatten“, die im Hintergrund einer Beziehung ihre Wirkung entfalten, wie z. B. der „Schatten der Zukunft“ im Falle der Reputation oder der „Schatten der Gerichte“, wenn solche zur Vertragsdurchsetzung ex post und kostenlos im Falle einer Beziehung zur Verfügung stehen.<sup>3</sup> Es handelt sich also um einen eher weitgesteckten Vertrauens-

<sup>2</sup> Vgl. dazu etwa die Beiträge in dem aktuellen Sammelband von Held / Kubon-Gilke / Sturm (2005) sowie die umfangreiche Monografie von Nooteboom (2002).

<sup>3</sup> Diese latenten Durchsetzungsmechanismen sind zu unterscheiden von solchen, die erst dann zur Verfügung stehen, wenn der vom Ausfall bedrohte Vertragspartner vor

begriff, für den Nooteboom auch den Überbegriff „reliance“ (s. u.) vorschlägt (ebenda, 49), wir sprechen in diesem Fall von „Vertrauen im weiteren Sinn“.

Dieser Vertrauensbegriff geht auf den amerikanischen Psychologen J. B. Rotter zurück und wird von aktuellen Beiträgen aus dem Bereich der experimentellen bzw. angewandten Spietheorie (Dufwenberg, 2002, 62), in denen ebenfalls strategische Investitionen thematisiert werden, verwendet: „Rotter defines (interpersonal) trust as an“ (ebenda): *„expectancy held by an individual that the word, promise, oral or written statement of individual or group can be relied on“* (Rotter, 1980, 1; Hervorhebung durch die Verfasser).

Genau diesen Begriff wollen wir in unserem Papier als *Vertrauen im weiteren Sinne* verwenden. Mit diesem Ansatz behaupten wir nicht, dass „reliance“ die einzige Möglichkeit ist, Vertrauen ökonomisch zu betrachten, sondern lediglich eine (wichtige) Variante. Jedoch eine Variante, die im Wirtschaftsalltag im Rahmen üblicher Rechtsgeschäfte eine zentrale Rolle spielt, die sich für die Modellierung strategischer Investitionen als zweckmäßig erwiesen hat und darüber hinaus den Vorteil besitzt, dass man sie auch mit den Methoden der formalen Vertragstheorie analysieren kann. Letztere bietet aber dem Ökonomen bekanntlich einen geeigneten Rahmen für die formale Betrachtung von Rechtsgeschäften und möglicherweise drohenden Rechtsbrüchen.

Nun fußt „trust on a non-calculative-basis“ auf der Erwartung, dass sich der Kooperationspartner in einem gewissen Sinn altruistisch verhalten wird. Viele mit der Verwendung des Vertrauensbegriffs anklingende Synonyme, Motivationen und Emotionen wie Sympathie, Loyalität, Aufrichtigkeit – sogar Liebe – verdeutlichen, dass dieser Vertrauensbegriff dem Alltagsempfinden weitaus näher kommt. Eine in der Sozialforschung etablierte Variante dieses „altruistischen“ Vertrauens ist „Trust-Responsiveness“. Unter Trust-Responsiveness versteht man die Disposition eines Akteurs Vertrauen zu erfüllen, wenn er merkt, dass es in ihn gesetzt wurde.<sup>4</sup> Der hier wirksam werdende Altruismus ist noch nicht erschöpfend erforscht, wird jedoch auf maßgeblich

---

oder im Laufe der Beziehung in sie investiert, also „aufrüstet“. Beispiele hierfür wären der Abschluß einer Versicherung, Einsatz von Monitoring-Technologien oder Hinterlegung eines Pfandes. Es herrscht in der Literatur Konsens darüber, dass in diesen Fällen nicht von Vertrauen gesprochen werden kann und sollte. Eine unter diesen Bedingungen erfolgende Vorleistung wäre nicht mehr „objektiv riskant“, da ein unbeteiligter Dritter in der Lage wäre, die Aufrüstungsaktivitäten zu beobachten und folglich zu erklären, warum die getätigte Vorleistung gerade kein besonderes Risiko darstellt.

<sup>4</sup> „Trust responsiveness is a tendency to fulfil trust because you believe that it has been placed on you, perhaps out of the feeling of not wanting to „let down“ the truster. A trust responsive agent is more likely to fulfil trust ... (the more she believes) that she is being trusted“ (Guerra/Zizzo, 2004, 25). Es ist nicht auszuschließen, dass auch der „warm glow“ Effekt – bekannt aus der Diskussion des spieltheoretisch fundierten Altruismusbegriffs – als Motivation dafür, empfangenes Vertrauen zu honorieren, experimentell einmal nachgewiesen werden kann.

zwei Motivationen zurückgeführt: Man möchte den anderen nicht enttäuschen, um entweder diesem einen Verlust zu ersparen (sog. „Let-Down-Aversion“) oder um zu verhindern, dass dieser seine Meinung über einen selbst (nachteilig) ändert.<sup>5</sup> Wir konzentrieren uns in diesem Beitrag auf die erste Form von „Trust-Responsiveness“, der „Let-Down-Aversion“, und bezeichnen sie im Folgenden mit *Vertrauen im engeren Sinn*.

In unserem Beitrag modellieren wir die Vertrauensbeziehung als spezifische, kooperative Investition (vgl. Schweizer, 2003, 2). Bereits das 1995 von Berg et al. in die Literatur eingeführte „investment game“ hat sich als außerordentlich ergiebig erwiesen, um Vertrauen im Verhalten von Individuen experimentell, also unter Laborbedingungen zu studieren (Cochard et al., 2004, 32 ff.). Phänomene wie „Reziprozität“, die Herausbildung von „Reputation“ u. ä. m. können mit Hilfe von Investitionsspielen vorzüglich beleuchtet werden.

Es wird die spezifische Investition des Vertrauensgebers allerdings nicht als mehrfach wiederholtes Spiel modelliert, sondern in einem institutionenökonomischen bzw. vertragstheoretischen Kontext. Dem Vertrauensgeber geht es dabei weniger um die moralischen Qualitäten des Vertrauensnehmers, auch nicht so sehr um dessen Kompetenz, sondern vielmehr darum, sich auf ihn verlassen zu können („reliance“).

### 3. Die „Waffel-Shop Story“

Das folgende Beispiel ist dem Beitrag von Cooter/Uhlen (1988) entnommen. Wir werden es zunächst im Sinne der Autoren reproduzieren, um es dann fortzuentwickeln. Dies wird uns erlauben, die Rolle von Vertrauen in Rechtsgeschäften und die möglichen Beziehung zwischen formellen Institutionen (hier: Gerichte und Schadensersatzregelungen) und Vertrauen genauer zu beleuchten.

Yvonne ist die Besitzerin eines Restaurants, genannt der „Waffel – Shop“. Seit dem Frühjahr bemerkt Yvonne bei sich bessernder wirtschaftlicher Lage einen deutlichen Nachfrageanstieg, den sie gerne durch Erwerb eines größeren Etablissements befriedigen möchte. Sie setzt sich daher mit der Baufirma von Xavier in Verbindung. Xavier verspricht (er ist der sogenannte „promisor“) Yvonne (sie ist die sogenannte „promisee“) bis zum 1. September (Neueröffnung des größeren Restaurants) den Bau fertig zu stellen. Xavier weiß, dass eine Vielzahl von Ereignissen vor dem 1. September eintreten könnten, die ihn an einer pünktlichen Fertigstellung des Gebäudes und der Inneneinrichtung hindern würden: So könnten Arbeiter streiken, ihm das Wetter einen bösen

<sup>5</sup> „Two elements in the informal explanations of trust responsiveness in the literature are aversion by the trustee to ‚letting down‘ the truster, and the idea that this aversion depends on the sympathy or respect the trustee feels for the trustor – on how ‚pro‘ his attitude towards her is.“ (Bacharach / Guerra / Zizzo, 2001, 6).

Streich spielen oder das Bauamt die Baugenehmigung in die Länge ziehen. Xavier kann aber, so nehmen wir an, die Wahrscheinlichkeit ( $p$ ) einer pünktlichen Fertigstellung des Neubaus selbst beeinflussen, indem er beispielsweise den Arbeitern einen Lohn über Tarif in Aussicht stellt, die Arbeitsorganisation so optimiert, dass an regenfreien Tagen über den Einbruch der Dunkelheit hinaus gearbeitet wird und schließlich das Bauamt von ihm „guten Zuspruch“ erfährt, das Genehmigungsverfahren zu beschleunigen. Alle diese Aktivitäten sind verschiedene Formen von „precaution“, also von Kosten verursachenden Vorsichtsmaßnahmen ( $x$ ), die aber dazu angetan sind, die Wahrscheinlichkeit zu erhöhen, dass Xavier sein Yvonne gegebenes Versprechen erfüllt. Soweit wir die Abmachung zwischen Xavier und Yvonne als Versprechen Xavier's auffassen liegt im Sinne von Richter/Furubotn eine Vertrauensbeziehung vor: Vertrauen im Sinne von Verlässlichkeit des Kooperationspartners.

#### 4. Betrachtung aus der Sicht der „Economics of Law“

Cooter/Ulen haben die Vertrauensbeziehung zwischen Yvonne und Xavier in einem Partialmodell der „Economics of Law“ modelliert, das mit den nicht-kooperativen Mengenstrategien traditioneller Duopolmodelle eng verwandt ist (vgl. Cooter/Ulen, 1988, 304 ff.): Jeder Spieler sucht sein eigenes Optimum, gegeben ein bestimmtes Verhalten des Partners. Für jeden Akteur gibt es zwei Modellbausteine; betrachten wir zunächst Xavier:

In Abbildung 1 (IV. Quadrant) wird der Zusammenhang zwischen der (stets positiven) Wahrscheinlichkeit  $p$  (dafür, dass der Neubau rechtzeitig fertig wird) und den Ausgaben  $x$  von Xavier zur Finanzierung von Vorsichtsmaßnahmen ( $X$ ) grafisch hergestellt. Wir erkennen einen positiven, aber degressiven Anstieg der Kurve, die sich asymptotisch der Horizontalen in Höhe von eins nähert. Es gilt:

$$p = p(x); p'(x) > 0; p''(x) < 0.$$

Der zweite Baustein für Xavier beschreibt den Zusammenhang zwischen den erwarteten gesamten Ausgaben ( $E$ ), die Xavier (über die normalen Baukosten hinaus, die Verfallstrafen) entstehen, in Abhängigkeit des Volumens der von ihm eingesetzten Vorsichtsmaßnahmen ( $X$ ). Wenn vereinfacht angenommen wird, dass eine Einheit von  $X$  einen US-Dollar kostet, dann sind die Ausgaben  $x$  mit der Menge der getroffenen Vorsichtsmaßnahmen  $X$  identisch; weitere Ausgaben kommen auf Xavier zu, wenn er sein Versprechen nicht halten kann; die Wahrscheinlichkeit dafür beträgt  $[1 - p(x)]$ . Wenn nämlich Yvonne ihr Vertrauen enttäuscht sieht, dann wird sie ein Gericht anrufen. Die Wahrscheinlichkeit dafür, dass Xavier den Rechtsstreit verliert, betrage  $q$  und die zu erwartende Schadenszahlung an Yvonne belaufe sich auf  $m$ . Die erwartete Geldbuße ( $D$ ) für Xavier entspricht also dem Produkt:  $D = qm$ .

Die wahrscheinlichen Prozessausgaben für Xavier ( $W$ ) hängen darüber hinaus von  $[1 - p(x)]$  ab:  $W = [1 - p(x)]qm = [1 - p(x)]D$ .

Somit erhalten wir als Summe der erwarteten Ausgaben für Xavier:  $E = x + [1 - p(x)]D$ .

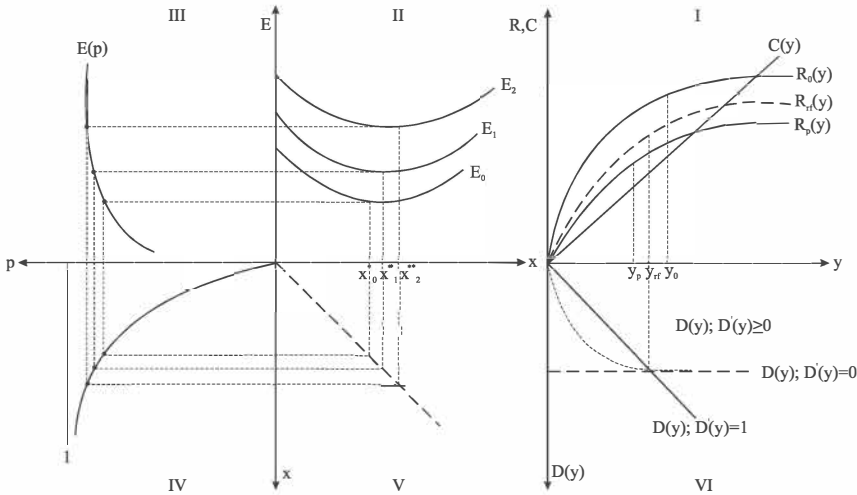


Abbildung 1

Je nachdem, ob Xavier es mit einer niedrigen ( $D_0$ ), einer mittleren ( $D_1$ ) oder einer hohen erwarteten Geldbuße ( $D_2$ ) zu tun hat, wird die nichtlineare Kurve in Abbildung 1 (II. Quadrant) weiter unten oder weiter oben verlaufen. Die Kurve ist deshalb nicht-linear, weil Xavier bei sehr geringem Einsatz von Vorsichtsmaßnahmen die wahrscheinlichen Ausgaben für die erwartete Geldbuße stark anhebt; erhöht er die Vorsichtsmaßnahmen, dann fallen dafür mehr Kosten an, die wahrscheinlichen Ausgaben für die erwartete Geldbuße sinken aber dafür. Solange der zweite Effekt überwiegt, sinken die erwarteten Ausgaben. Allerdings wird die Kurve dann wieder ansteigen, sobald die erhöhten Ausgaben für Vorsichtsmaßnahmen die verbleibenden wahrscheinlichen Ausgaben für die erwartete Geldbuße übersteigen. Der optimale Einsatz von Vorsichtsmaßnahmen ist für Xavier dort erreicht, wo die Grenzkosten für die Erhöhung von Vorsichtsmaßnahmen gerade dem Grenznutzen durch die Reduktion der wahrscheinlichen Ausgaben für die erwartete Geldbuße entsprechen, in der Grafik also bei den jeweiligen  $x^*$  ( $x_0^*$ ,  $x_1^*$ ,  $x_2^*$ ).

Das Partialkalkül von Yvonne sieht hingegen wie folgt aus: Sie verlässt sich („reliance“) zunächst einmal auf die Zusage von Xavier, den Neubau des Restaurants termingerecht fertig zu stellen. Um ihre zusätzlichen Gäste bereits am 1. September bewirten zu können, wird sie zusätzliches Personal einstellen.



Die Kosten ( $C$ ) hierfür seien direkt proportional zu dem Ausmaß, in dem sie sich auf Xavier verlässt ( $y$ ):  $C = h(y)$ ;  $h' > 0$ ;  $h'' = 0$ .

Der Betrag der Geldentschädigung, den Yvonne erhält, wenn es zum Prozeß kommt, ergibt sich aus ihrer Sicht als:  $D = D(y)$ ;  $D'(y) \geq 0$ .

Je mehr sie sich nämlich auf Xavier verlässt, desto größer werden die Ausgaben für zusätzliches Personal ausfallen und demzufolge ihr Verlust betragen. Ein größerer Verlust kann allerdings per Gerichtsentscheid durch eine höhere Geldentschädigung aufgewogen werden (Abbildung 1, VI. Quadrant).

Die Entschädigungsfunktion kann allerdings unterschiedliche Verläufe annehmen. Bei einer monoton ansteigenden Funktion würde das Gericht Yvonne die vollen Aufwendungen für die Einstellung des zusätzlichen Personals zugestehen. Je mehr Personal Yvonne einstellt, desto höher fällt die zu erwartende Entschädigung aus. Eine solche Entschädigungsregel ist allerdings mit der Versuchung für Yvonne verbunden, besonders viel Personal anzuheuern und sich besonders stark auf Xavier zu verlassen; genauer: Das Risiko, sich auf Xavier zu verlassen, wird dabei von Yvonne vollständig an Xavier zurückübertragen. Eine vollständige Risikozurückübertragung ist genau dann gegeben, wenn gilt:  $D'(y) = 1$ .

Die zwei in der Rechtspraxis etabliertesten Schadensersatzregeln besitzen diese Eigenschaft: Der Ersatz des Schadens, der durch Nichterfüllung zustande kommt (*expectation damages*) und der Ersatz des Vertrauensschadens (*reliance damages*).

Um eine solche, das Eingehen von Risiken zu stark fördernde Anreizstruktur zu beseitigen, muss die erwartete Entschädigungszahlung von dem Ausmaß, mit dem sich Yvonne mit Zusatzpersonal eindeckt, also auf Xavier vertraut, entkoppelt werden:  $D'(y) = 0$ .

Jetzt trägt Yvonne das Risiko einer weiteren Einheit „Personal“ bzw. der „reliance“ darauf vollständig. Auch das erscheint eine sehr spezielle Lösung. Gäbe es nicht eine (dritte) Regel, welche das Risiko zwischen Vertrauensgeber und -nehmer effizient verteilt? Effizient wäre eine Lösung dann, wenn sie dazu führt, dass eine marginale Zunahme in der Geldbuße für Xavier gerade gleich groß ist dem voraussehbaren Zusatzvorteil (Grenzerlös minus Grenzkosten), den Yvonne bei einem Bezug des neuen Restaurants mit zusätzlichen Gästen und erweitertem Personal erzielt. Voraussehbar heißt, einen Wert *zwischen* dem allzu optimistischen und dem allzu pessimistischen Erlösszenario als plausibel anzunehmen, im Englischen: *„reasonably foreseeable“* (*rf*):  $D_{rf} = D(y_{rf})$ . Man spricht in diesem Zusammenhang auch von „individuell geregelten Schadensersatzfällen“ (*liquidated damages*).

Hierauf kommen wir weiter unten noch einmal zurück. Der zweite Baustein von Yvannes Kalkül beinhaltet die folgende Überlegung: Neben den oben geschilderten Kosten erwartet sie (Netto-)Einnahmen ( $R$ ) ab dem 1. September



aus dem Restaurantbetrieb. Dafür gibt es allerdings sowohl ein optimistisches Szenario ( $R_o$ ) – Xavier wird rechtzeitig fertig und mehr Gäste können bewirtet werden – und ein pessimistisches Szenario ( $R_p$ ) – Xavier wird nicht fertig, das alte, wesentlich kleinere Restaurant muss weiter genutzt werden. In beiden Szenarien ist der Verlauf der Brutto- bzw. Nettoeinnahmen eine Funktion vom Ausmaß, mit dem Yvonne sich auf Xavier verlässt:

$$R_p = R_p(y); R'_p(y) > 0; R''_p(y) < 0 \quad \text{und} \quad R_o = R_o(y); R'_o(y) > 0; R''_o(y) < 0 .$$

Das jeweilige Partialoptimum für Yvonne liegt nun dort, wo ihre Grenzkosten den erwarteten Grenzerlösen entsprechen. Dies sind in Abbildung 1 (I. Quadrant) jeweils die Werte  $y_o$  und  $y_p$ .

In den Quadranten I und VI der Abbildung 1 lassen sich die beiden Partialkalküle von Xavier und von Yvonne zusammenführen und man kann zusätzlich den Versuch unternehmen, die oben abgeleitete Effizienzregel anzuwenden:

$$R'_{rf}(y_{rf}) - C'_{rf}(y_{rf}) = D'_{rf}(y_{rf}) = 0 .$$

Diese Regel wurde im sechsten und ersten Quadranten von Abbildung 1 realisiert, weil, wie man sieht, in Höhe der realistischen reliance ( $y_{rf}$ ) der Grenzscha-den gleich null ist und zugleich Grenzerlös und Grenzkosten für Yvonne sich entsprechen, ihre Differenz mithin null ist. Zugleich haben wir durch Kombination von Quadrant II aus Abbildung 1 mit einer 45-Grad-Linie (im V. Quadranten) den impliziten Zusammenhang zwischen den Ausgaben von Xavier ( $E$ ) einerseits und der Wahrscheinlichkeit dafür, dass er sein Versprechen erfüllen kann ( $p$ ), dargestellt (III. Quadrant in Abbildung 1).

## 5. Relevante Referenzlösungen der formalen Vertragstheorie

Im Zusammenhang mit der oben dargestellten Situation stellt sich nun die Frage, wie die optimale Zusammenarbeit von Yvonne und Xavier ausgestaltet sein sollte, insbesondere mit welcher Intensität Yvonne sich bereits vor ihrer Entscheidung zur Neueinstellung von Personal vertraglich gegen eine mögliche Nichterfüllung von Xavier absichern sollte. Um dies beurteilen zu können, ist es zunächst erforderlich, einige Referenzlösungen vorzustellen: In einem ersten Schritt wird die sozial optimale Lösung (first-best) hergeleitet, die als „Wert der Beziehung“ (zwischen Xavier und Yvonne) interpretierbar ist. Ohne jegliche vertragliche Absicherung oder sonstige Erweiterungen des Kontexts (z. B. um Reputationseffekte, Triggerstrategien etc.) wird das größere Restaurant von Xavier mangels Anreiz nicht gebaut und Yvonne wird dies auch sicher antizipieren (Inaktivität). Es stellt sich mithin das bekannte Ergebnis eines „trust games“ (Kreps / Milgrom et al., 1982, 245 f.) ein. Dieses Dilemma lässt sich, wie wir sehen werden, auch nicht mithilfe eines anreiz-

kompatiblen Vertrages zwischen Yvonne und Xavier lösen, denn im Rahmen einer solchen leistungsabhängigen Regelung erhielte Xavier einen zu großen Anteil am Zugewinn, d. h. Yvonne würde Verluste machen. Anschließend wird gezeigt, dass die drei bekannten Schadensersatzregeln die notwendigen Anreize zur Zusammenarbeit schaffen, wenngleich nur eine von ihnen gesamtwirtschaftlich effizient ist (first-best).

Zusätzlich zu den bisher getroffenen Annahmen für Yvones Erlösfunktion  $R(y)$  gelte für ein beliebiges  $\tilde{y}$ :  $R_o(\tilde{y}) > R_p(\tilde{y})$  sowie  $R'_o(\tilde{y}) > R'_p(\tilde{y})$ .

Yvones erwarteter Nettogewinn<sup>6</sup> für den Monat September beträgt dann:

$$E(\pi_y) = p(x)R_o(y) + [1 - p(x)]R_p(y) - y - [R_p(y_p) - y_p] - P.$$

Der erste Teil auf der rechten Seite des Gleichheitszeichens bis zur zweiten eckigen Klammer ist ihr erwarteter Bruttogewinn aus der Zusammenarbeit mit Xavier. Der dann folgende Term in der eckigen Klammer ist ihre Außenoption, also der sichere Gewinn, den sie erhält, wenn sie Xavier nicht vertraut und sich entsprechend auch nicht auf die Zusammenarbeit einlässt.<sup>7</sup> Der Preis  $P$  ist die Bezahlung für Xavier, und zwar dafür, dass er überhaupt das größere Restaurant baut, nicht dafür, dass der Neubau pünktlich zum 1. September bezogen werden kann.<sup>8</sup>

Das Optimierungsproblem von Yvonne lautet damit:  $\max_y E(\pi_y)$ .

Als Bedingung erster Ordnung erhält man:  $R'_p(y) + p(x)[R'_o(y) - R'_p(y)] = 1$ .

Das Wertgrenzprodukt im alten Restaurant zusammen mit dem erwarteten Zuwachs des Wertgrenzprodukts durch die Restauranterweiterung (linke Seite) muß ihren Grenzkosten (rechte Seite) entsprechen. Würde sie Xavier voll vertrauen oder die Restauranterweiterung vollkommen unabhängig von Xavier durchführen können ( $p = 1$ : „absolutes“ Vertrauen oder autonome Produktion), dann könnte sie sich den vollen Zuwachs des Wertgrenzprodukts  $[R'_o(y) - R'_p(y)]$  sichern. In diesem Fall absoluter Sicherheit kann sie auch ihre Investition risikolos vornehmen und  $y$  entsprechend  $R'_o(y) = 1$  optimal an die Kapazitäten des (als sicher zu erwartenden) größeren Restaurants anpassen:

<sup>6</sup> „Nettogewinn“ meint den Gewinn in Bezug zur Außenoption, also in Bezug zur Referenzsituation im kleineren Restaurant mit weniger Personal.

<sup>7</sup> Sie bleibt also im September in ihrem kleineren Restaurant mit der hierfür optimalen Anzahl an Kellnern und Küchenkräften  $y_p$  (status quo). Dies impliziert, dass sie sich nicht kurzfristig an eine andere Baufirma wenden kann.

<sup>8</sup>  $P$  ist ein Fixum und damit für Xavier nicht entscheidungsrelevant. Aus  $P$  entstehen also keine Anreize für Xavier, den Neubau termingerecht fertigzustellen, worum es Yvonne bei ihrer Entscheidung geht. Weiterhin ist die Höhe von  $P$  für die Bestimmung des sozialen Optimums irrelevant, denn eine Variation von  $P$  induziert lediglich eine Umverteilung des sozialen Überschusses zwischen beiden Parteien nachdem sie ihre Wahl getroffen haben.

$y^* = y_o$ . Im entgegengesetzten Extremfall des absoluten Misstrauens ( $p = 0$ ) geht sie gedanklich sicher davon aus, dass sie am 1. September noch immer im kleinen Restaurant sitzt, entsprechend löst sie das Optimierungsproblem durch  $R'_p(y) = 1$ . In diesem Szenario ist für sie jeder noch so kleine Investitionsbetrag in die Restauranterweiterung verlorenes Geld:  $y^* = y_p$ . Aus diesen Überlegungen wird deutlich: Yvones Einschätzung („Vertrauen“) darüber, inwieweit Xavier sich an sein Versprechen halten wird, bestimmt das optimale Niveau der von ihr im voraus getätigten und daher risikobehafteten Investition  $y$ :  $y^* = y(p)$ .

Die Gewinnfunktion ohne Schadensersatzzahlungen (d. h.  $D = 0$ ) von Xavier lautet:  $\pi_x = P - x$ . Zur Bestimmung des sozialen Überschusses wird der gemeinsame (erwartete) Gewinn maximiert:

$$\max_{x,y} E(\pi_s) = E(\pi_y) + \pi_x = p(x)R_o(y) + [1 - p(x)]R_p(y) - y - [R_p(y_p) - y_p] - x.$$

Die beiden Bedingungen erster Ordnung für ein Maximum lauten:

$$\frac{\partial E(\pi)}{\partial y} = R'_p(y) + p(x)[R'_o(y) - R'_p(y)] = 1$$

$$\frac{\partial E(\pi)}{\partial x} = p'(x)[R_o(y) - R_p(y)] = 1.$$

Setzt man die untere in die obere Gleichung ein, dann muss im sozialen Optimum gelten:

$$R'_p(y) + p(x)[R'_o(y) - R'_p(y)] = p'(x)[R_o(y) - R_p(y)].$$

Die linke Seite ist die bekannte Effizienzbedingung für Yvones Optimierungsproblem: Je stärker sie auf die pünktliche Neueröffnung des Restaurants vertraut, desto eher wird sie für die Produktionsplanung die höhere Erlösfunktion  $R_o(y)$  zugrunde legen (Abbildung 2). Gemäß Annahme ist dies c.p. gleichbedeutend mit einem Anstieg des Grenzerlöses, d. h. die gewinnmaximale Produktion erfordert nun einen höheren Einsatz an Produktionsfaktoren im Vergleich zur Situation des kleinen Restaurants. Ihr Beitrag zum sozialen Überschuss besteht also darin, dass sie – für eine gegebene Wahrscheinlichkeit  $p$  – ihr Produktionsproblem optimal anpasst („Auslastungseffekt“: Bewegung von Punkt 2 zu Punkt 3 in Abbildung 2). Die rechte Seite bildet Xaviers marginalen Beitrag zum sozialen Überschuss ab. Dabei ist der Term in der eckigen Klammer der Erlöszuwachs für Yvonne, wenn Xavier ihr pünktlich das größere Restaurant hinstellt. Wie hoch dieser Zuwachs überhaupt ausfallen kann, hängt von Yvones Entscheidung ab: Hat sie „vertrauensvoll“ zusätzliches Personal eingestellt, so kann sie den Kapazitätswachstum gewinnbringend ausnutzen (Bewegung von Punkt 4 zu Punkt 3 in Abbildung 2), hat

sie hingegen von zusätzlichen Einstellungen abgesehen, muss sie nun mit ihrem bisherigen Personal, also beispielsweise mit ihrem einzigen Kellner und ihrer einzigen Küchenkraft, das größere Restaurant betreiben. Dadurch wird zwar auch ein Erlöszuwachs realisiert (Bewegung von Punkt 1 zu Punkt 2), jedoch ein kleinerer. Xaviers Beitrag zum sozialen Überschuss besteht also in der Bereitstellung der zusätzlichen Kapazität („Kapazitätseffekt“) – inwieweit diese höhere Kapazität allerdings auch wirklich ausgeschöpft wird, hängt von Yvonne ab. Der Faktor  $p'(x)$  vor der eckigen Klammer lässt sich anschaulich als „Wirkungsgrad“ von Xaviers Anstrengung auffassen: Er beschreibt, in welchem Umfang es Xavier gelingt, durch marginale Erhöhung seines Anstrengungsniveaus die Wahrscheinlichkeit für die pünktliche Fertigstellung des Neubaus zu erhöhen.<sup>9</sup> Dies verdeutlicht eine notwendige Bedingung für ein effizient in andere Personen gesetztes Vertrauen: Es sollte nur maximal in dem Maße vertraut werden, wie der Vertrauensnehmer (hier: Xavier) auf das vom Vertrauensgeber (hier: Yvonne) erwartete Ergebnis effektiv Einfluss nehmen kann.<sup>10</sup> Die Bedingungen 1. Ordnung von Yvonne bzw. Xavier entsprechen grafisch ihren jeweiligen Reaktionskurven (Abbildung 3).<sup>11</sup>

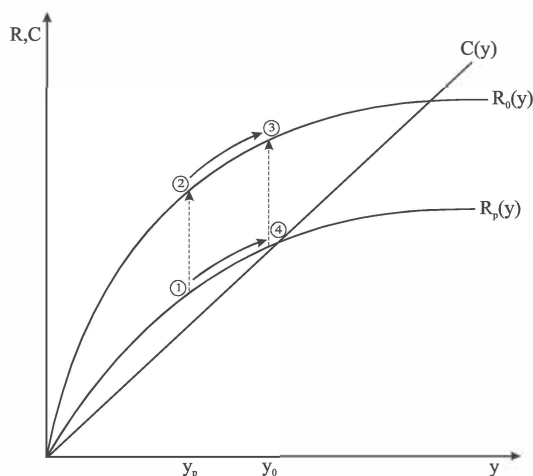


Abbildung 2

<sup>9</sup> Vgl. dazu auch den Verlauf der Funktion  $p(x)$  im IV. Quadranten in Abbildung 1.

<sup>10</sup> Diese Bedingung bezieht sich also weniger auf seine Motivation, sondern auf seine technologischen Grenzen. In diesem Sinne lässt sich die Funktion  $p(x)$  (aus Quadrant IV, Abb. 1) auch durchaus als „Produktionsfunktion der Vertrauensbefriedigung“ auffassen: Je niedriger die Grenzproduktivität, desto teurer wird es für Xavier, die Erwartungen von Yvonne zu bestätigen.

<sup>11</sup> Das Krümmungsverhalten der Reaktionskurven (Voraussetzung für ein inneres Optimum) ergibt sich direkt aus allen bisher gemachten Annahmen: Die Reaktionskurven  $x^*(y)$  für Xavier sowie  $y^*(x)$  für Yvonne müssen demnach konkav sein.

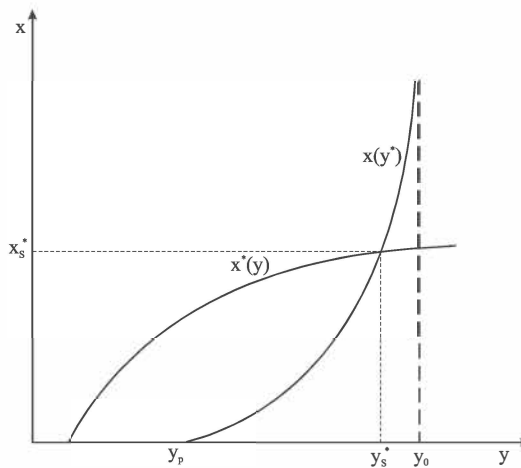


Abbildung 3

Der Schnittpunkt beider Kurven  $(x_s^*, y_s^*)$  erfüllt die oben hergeleitete Effizienzbedingung und maximiert den sozialen Überschuss. Man erkennt, dass das sozial optimale Niveau von  $y(y_s^*)$  kleiner ist als dasjenige bei autonomer Produktion ( $y_0$ ). Dies liegt daran, dass bei der Bestimmung des sozial optimalen Vertrauenslevels auch den Anstrengungskosten von Xavier Rechnung getragen werden muss. Damit ist das soziale Optimum beschrieben.

Es stellt sich nun die Frage, unter welchen Bedingungen Yvonne und Xavier tatsächlich auch einen Anreiz dafür haben,  $y$  bzw.  $x$  in der jeweils effizienten Höhe zu wählen! Geht man von den obigen, isolierten Zielfunktionen aus, dann wählt Xavier trivialerweise  $x^* = 0$ , denn seinen Grenzkosten in Höhe von 1 stehen Grenzerlöse in Höhe von 0 gegenüber. Ohne zusätzliche Anreize wird sich Xavier nicht anstrengen. Dies antizipiert Yvonne ( $p = 0$ ) und wählt  $y^* = y_p$ . Wenn sicher antizipiert werden kann, dass gegebenes Vertrauen enttäuscht werden wird, wird es gar nicht erst geschenkt, somit erreichen beide (Yvonne und Xavier) lediglich die Außenoption (Inaktivität) und der Wert der Beziehung beträgt 0. Diese Situation stellt ein (soziales) Dilemma dar. Dieses Dilemma lässt sich auch nicht durch einen anreizkompatiblen Vertrag bewältigen, den Yvonne Xavier u. U. unterbreitet. Die Grundidee eines solchen Vertrages ist es, eine variable Bonuszahlung  $B(x)$  in Abhängigkeit des geleisteten Anstrengungsniveaus zu vereinbaren.<sup>12</sup> Dabei sollte die Bonuszahlung so gewählt werden, dass Xavier automatisch zur sozial optimalen Wahl  $x_s^*$  veranlasst wird. Entsprechend seiner Bedingung 1. Ordnung wählt Xavier  $x_s^*$  für  $p'(x_s^*)[R_o(y_s^*) - R_p(y_s^*)] = 1$ , d. h. dann, wenn Yvonne auf dem effizienten

<sup>12</sup> Dies ist möglich, weil die Variable  $x$  ex-post verifizierbar ist. Darin unterscheidet sich das vorliegende Problem von den Lösungsansätzen des klassischen Moral-Hazard.

Niveau vertraut und der dadurch ermöglichte Erlöszuwachs in voller Höhe von Yvonne an Xavier übergeht. Xavier schöpft somit den gesamten Kapazitätseffekt ab. In diesem Fall bliebe Yvonne lediglich der Auslastungseffekt (Punkt 4 in Abbildung 2) für das alte, bekanntlich kleinere (!) Restaurant:

$$\pi_y = [R_p(y_s^*) - R_p(y_p) - (y_s^* - y_p)] - P < 0 .$$

Nachdem sie Xavier ausgezahlt hat steht sie so da, als hätte sie die größere Personalausstattung für das kleinere Restaurant gewählt, d. h. sie erleidet einen Nettoverlust (Term in eckigen Klammern). Das aber bedeutet, dass es für Yvonne auch keinen Sinn mehr machen würde, den Preis  $P$  im Gegenzug herabzusetzen. Da bereits der Term in der eckigen Klammer negativ ist, ist der Deal mit Xavier unter dem Strich selbst dann noch für sie defizitär, wenn Xavier ihr das neue Restaurant kostenlos bauen würde ( $P = 0$ ). Daher ist ein anreizkompatibler Kontrakt in diesem Fall nicht implementierbar. Wie aus der Diskussion des vorigen Kapitels deutlich wurde, ist die Einrichtung einer Schadensersatz-Klausel grundsätzlich dazu geeignet, das beschriebene Dilemma zu überwinden. Die Literatur unterscheidet hier insbesondere zwischen dem Ersatz des Nichterfüllungsschadens („expectation damages“  $D_E$ )<sup>13</sup>, dem Ersatz des Vertrauensschadens („reliance damages“  $D_R$ )<sup>14</sup> sowie den sog. individuell geregelten Schadensersatzfällen („liquidated damages“  $D_L$ ).

Beim Ersatz des Nichterfüllungsschadens wird Yvonne im Schadensfall so gestellt, als hätte Xavier ihr das größere Restaurant gebaut und sie somit den maximalen Gewinn hätte realisieren können, d. h.  $D_E = R_o(y) - R_p(y)$ . Der erwartete Nettogewinn von Yvonne für den Monat September beliefe sich damit auf:

$$E(\pi_y) = p(x)R_o(y) + [1 - p(x)][R_p(y) + D_E] - y - [R_p(y_p) - y_p] - P$$

und vereinfacht sich nach Einsetzen von  $D_E$  und Umformen zu

$$\pi_y = R_o(y) - R_p(y_p) - (y - y_p) - P .$$

Der maximal erreichbare Nettogewinn ist ihr nun sicher, d. h. er ist unabhängig von der Wahrscheinlichkeit  $p(x)$ . Dies ist der Fall der vollständigen Risikozurückübertragung von Yvonne an Xavier, d. h. Yvonne ist vollversichert.<sup>15</sup>

<sup>13</sup> Der deutsche Gesetzgeber spricht verschiedentlich von „Schadensersatz statt der Leistung“ (§§ 281 ff., 311a Abs. 2 BGB) bzw. von „Schadensersatz wegen Nichterfüllung“ (§ 651 f. Abs. 1 BGB). Allgemeiner wird in diesem Zusammenhang vom „positiven Interesse“ gesprochen.

<sup>14</sup> Gesetzliche Anwendungsfälle für den Vertrauensschaden finden sich bspw. in §§ 119, §§ 120, § 122, § 179 Abs. 2 BGB. Allgemeiner wird in diesem Zusammenhang auch vom „negativen Interesse“ gesprochen.

Wie beim Fall des absoluten Vertrauens bzw. der autonomen Produktion löst sie  $R'_o(y) = 1$  und wählt  $y^* = y_{DE} = y_o$ .

Xaviers erwarteter Gewinn beträgt bei Ersatz des Nichterfüllungsschadens:

$$E(\pi_x) = P - x - [1 - p(x)] D_E = P - x - [1 - p(x)] [R_o(y) - R_p(y)] .$$

Die Bedingung 1. Ordnung liefert:

$$E'(\pi_x) = p'(x) [R_o(y) - R_p(y)] = 1 .$$

Das ist dieselbe Marginalbedingung (und damit auch die identische Reaktionskurve) wie bei der Herleitung des sozialen Optimums, d. h. Xavier wählt  $x$  stets auf dem sozial effizienten Niveau  $x_{DE} = x_s^*$ . Lediglich Yvonne vertraut zu stark, denn  $y_{DE} > y_s^*$ . Dieses Ergebnis ist – wie von Shavell gezeigt wurde – allgemeingültig für die Schadensersatzklausel bei Nichterfüllung im Kontext spezifischer Investitionen (vgl. Shavell, 1980, 466 f.). In Abbildung 4 ist dieser Fall im Vergleich zur sozial optimalen Lösung eingetragen.

Beim Ersatz des Vertrauensschadens wird Yvonne im Schadensfall so gestellt, als hätte es die Zusammenarbeit mit Xavier (bzw. den Vertrag) nie gegeben, d. h. ihr wird die Außenoption garantiert. Der Schadensersatz beträgt dann  $D_R = (y - y_p) - [R_p(y) - R_p(y_p)]$ . Der erwartete Nettogewinn von Yvonne für den Monat September beläuft sich damit auf:

$$E(\pi_y) = p(x) R_o(y) + [1 - p(x)] [R_p(y) + D_R - y - [R_p(y_p) - y_p]] - P$$

und vereinfacht sich nach Einsetzen von  $D_R$  und Umformen zu

$$E(\pi_y) = p(x) [R_o(y) - R_p(y_p) - (y - y_p)] .$$

Bricht Xavier sein Versprechen ( $p = 0$ ), dann beträgt Yvonnens Nettogewinn 0, d. h. sie ist in diesem Fall zwar nicht vollversichert, jedoch gegen Verluste immunisiert. Erfüllt er dagegen sein Versprechen, dann kann sie einen umso höheren Gewinn realisieren, je stärker sie ihm vertraut hat. Vertrauen ist also nach wie vor ihre (schwach) dominante Strategie, wie sich auch an der Bedingung 1. Ordnung zeigt: Wiederum hat sie im Optimum  $R'_o(y) = 1$  (absolutes Vertrauen bzw. autonome Produktion) zu erfüllen und wählt  $y^* = y_{DR} = y_o$ . Also auch in dem Fall, wenn ihr nur der Vertrauensschaden ersetzt wird, hat sie einen Anreiz, zu stark zu vertrauen. Xaviers erwarteter Gewinn beträgt bei Ersatz des Vertrauensschadens:

$$E(\pi_x) = P - x - [1 - p(x)] D_R = P - x - [1 - p(x)] [y - y_p - R_p(y) + R_p(y_p)] .$$

<sup>15</sup> Die Schadensersatzklausel bei Nichterfüllung ist insofern ein klassisches Beispiel für Moral-Hazard seitens des Prinzipals.



Die Bedingung 1. Ordnung liefert jetzt:

$$E'(\pi_x) = p'(x)[y - y_p - R_p(y) + R_p(y_p)] = 1.$$

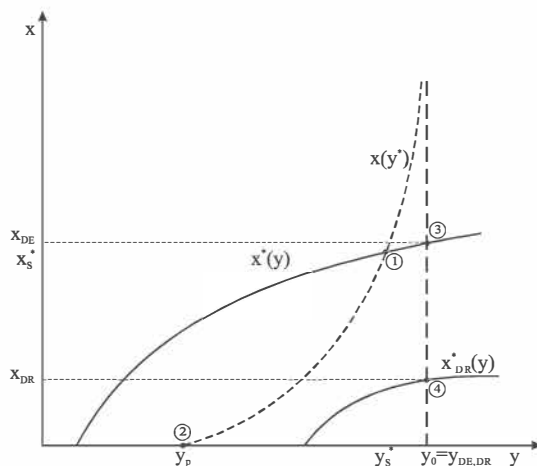


Abbildung 4

Xavier haftet nur für die Auslagen von Yvonne, die sie im Vertrauen darauf tätigte, dass Xavier seine Zusage erfüllt. Konkret wäre dies der Monatslohn des zusätzlich für September eingestellten Personals abzüglich des mit diesem Personal zusätzlich erwirtschafteten Erlöses. Daher wird der Betrag in den eckigen Klammern deutlich niedriger sein als im Fall des sozialen Optimums oder des Schadensersatzes bei Nichterfüllung. Das heißt aber nichts anderes, als dass sein Grenzerlös in diesem Szenario signifikant kleiner ausfallen wird als in den angesprochenen zwei Vergleichsfällen. Für ein gegebenes Vertrauensniveau  $y$  wird Xavier also ein signifikant niedrigeres  $x_{DR}$  wählen. Xavier wird also  $x$  nicht mehr auf dem sozial effizienten Niveau wählen, entsprechend verschiebt sich seine Reaktionsfunktion nach unten (vgl. Abbildung 4). Die Situation beim Vertrauensschadensersatz ist also einerseits gekennzeichnet durch zu starkes Vertrauen (ineffizientes Niveau  $y_{DR} > y_s^*$ ) und andererseits durch ein deutlich zu niedriges Anstrengungsniveau (ineffizientes Niveau  $x_{DR} < x_s^*$ ).<sup>16</sup>

Der letzte Fall betrachtet die sog. individuell geregelten Schadensersatzfälle „liquidated damages“ ( $D_L$ ), deren Eigenschaften bereits im vorigen Kapitel

<sup>16</sup> Allgemein lässt sich zeigen, dass der Ersatz des Nichterfüllungsschadens den Fall des Vertrauensschadensersatzes pareto-dominiert, d. h. für jeden Fall mit Vertrauensschadens-Klausel lässt sich ein mindestens ebenso effizientes Ergebnis mithilfe einer Nichterfüllungsschadens-Klausel erzielen. (Vgl. Rogerson, 1984, 40 sowie Che / Chung, 1999, 87).

untersucht wurden. Wie die Argumentation zeigte, ist diese Art des Schadensersatzes geeignet, die sozial optimale Lösung herbeizuführen. Die Schadensersatzzahlung muss hierfür unabhängig sein von  $y$  (um den Moral-Hazard-Effekt bei Yvonne auszuschalten), gleichzeitig muss  $\overline{D}_L$  auf die Höhe gesetzt werden, bei der Vertrauen effizient ist. Gemäß seiner Reaktionskurve wird Xavier dann das effiziente Anstrengungsniveau  $x_s^*$  wählen, wenn Yvonne auf dem effizienten Niveau  $y_s^*$  vertraut:

$$p'(x)[R_o(y_s^*) - R_p(y_s^*)] = 1.$$

Hieraus ergibt sich unmittelbar die festzulegende Höhe des Schadensersatzes:

$$\overline{D}_L = R_o(y_s^*) - R_p(y_s^*).$$

Auf diese Weise lässt sich zwar die erstbeste Lösung erreichen, ihre Implementierung in der Praxis stellt sich jedoch problematisch dar: Um die optimale Schadenshöhe bestimmen zu können, müsste beiden  $y_s^*$  bekannt sein. Näherungsweise kann hierfür  $y_{\mathcal{H}}$  (vgl. Kapitel 3) herangezogen werden, d. h. beide müssten sich auf einen Abschlag von Yvannes privater, gewinnmaximaler Wahl  $y_o$  einigen. Um zu zeigen, unter welchen Bedingungen eine solche Lösung zustande kommt, müsste explizit der vorgeschaltete Verhandlungsprozess zwischen Xavier und Yvonne modelliert werden, worauf wir hier aber verzichten.

## 6. Imperfekte Jurisdiktion

Bisher wurde implizit unterstellt, dass das Gericht stets eine perfekte Entscheidung trifft ( $q = 1$ ), d. h. Xavier kann sicher damit rechnen, im Falle eines (für Yvonne) zu niedrigen Anstrengungsniveaus auf die Leistung von Schadensersatz verklagt zu werden. Diese Annahme ist aus mehreren Gründen unrealistisch. Erstens liegt eine Situation asymmetrischer Information vor, d. h. das Gericht ist über die ursprüngliche Vereinbarung sowie über die Umstände, die zum Vertragsbruch führten, schlechter informiert als die Parteien Yvonne und Xavier. Zwar wird immer eine von beiden Parteien einen Anreiz haben, seine private Information dem Gericht gegenüber zu offenbaren, jedoch hilft dieser Umstand dem Gericht wenig: Die andere Partei wird eine Gegendarstellung liefern, es stünde Aussage gegen Aussage. Zweitens bleibt einem Gericht, selbst bei korrekter Darstellung des Sachverhalts, oft noch sehr viel Interpretationsspielraum. Etwa die Frage, ob Yvonne tatsächlich im größeren Restaurant den Umsatz  $R_o(y_o)$  erzielt hätte, muss dem Gericht zunächst mithilfe aufwendiger, teurer Gutachten nachgewiesen werden. Schließlich sind auch Fälle denkbar, in denen ein Gericht überhaupt das Zustandekommen des Vertrages bestreitet, z. B. weil Yvonne und Xavier bestimmte juristische Vorkehrungen (Formvorschriften etc.) nicht getroffen bzw. erfüllt haben. Yvonne

wird also nie mit absoluter Sicherheit davon ausgehen können, dass sie im Schadensfall entschädigt wird und für Xavier wirkt das Gericht weniger abschreckend, denn es besteht zumindest eine kleine Wahrscheinlichkeit dafür, dass Yvones Klage erfolglos bleibt. Welche Konsequenzen hat ein solches „weiches“ Gericht für die sozial optimale Lösung? Ausgangspunkt der Betrachtung ist der Fall des Nichterfüllungsschadens ( $D_E$ ), die folgenden Überlegungen gelten aber in äquivalenter Weise auch für den Ersatz des Vertrauensschadens ( $D_R$ ).

Die im vorigen Kapitel hergeleitete (zweitbeste) Lösung war gekennzeichnet von einem effizienten Anstrengungsniveau ( $x_{DE} = x_s^*$ ) und zu starkem Vertrauen ( $y_o > y_s^*$ ). Analog zum Lösungsansatz aus Kapitel 5 lauten die Bedingungen erster Ordnung bei expliziter Berücksichtigung des (i. d. R. unvollkommenen) Perfektionsgrades des Gerichts ( $q$ ) folgendermaßen<sup>17</sup>:

$$E'(\pi_Y) = p[R'_o(y) - R'_p] + R'_p(y) + q[1 - p(x)][1 - R'_p(y)] = 1$$

$$E'(\pi_x) = qp'(x)[R_o(y) - R_p(y)] = 1.$$

Für  $q = 1$  erhält man die optimale Lösung für den Ersatz des Nichterfüllungsschadens. Reduziert man  $q$  schrittweise ausgehend von  $(x_{DE}/y_{DE})$ , dann reduziert sich bei Yvonne der Grenzertrag des Vertrauens um den Betrag  $[1 - p(x)][1 - R'_p(y)]$ , c.p. sinkt  $y^*(x)$ . Es gilt:  $\partial y^*/\partial q \geq 0$ .

Yvones Reaktionskurve geht mit sinkendem  $q$  von der Senkrechten in  $y_o$  immer mehr zu ihrer Indifferenzkurve der first-best-Lösung über (vgl. gestrichelte Linien in Abbildung 5). Dies lässt sich auch formal zeigen: Setzt man in obiger Bedingung erster Ordnung  $q = 0$ , dann erhält man die bekannte Bedingung erster Ordnung für die Maximierung des sozialen Überschusses (first-best) aus Kapitel 5. Die Feststellung lässt sich folgendermaßen interpretieren: Wenn sich Yvonne der Anstrengung von Xavier sicher sein könnte, dann verursacht die Existenz des Gerichts einen Effizienzverlust, da die Schadensersatzklausel einen Versicherungseffekt in sich trägt. Bei isolierter Betrachtung von ausschließlich Yvones Entscheidung wäre  $q = 0$  die effiziente Institution (also kein Gericht und kein Schadensersatz, wodurch ihr Moral Hazard ausgeschaltet wird).

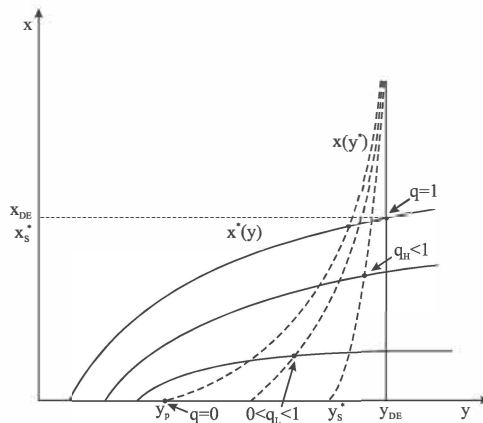
Bei Reduktion von  $q$  sinkt der Grenzertrag der Anstrengung für Xavier um den Betrag  $p'(x)[R_o(y) - R_p(y)]$ , c.p. sinkt auch  $x^*(y)$ . Es gilt:  $\partial x^*/\partial q \geq 0$ .

Wenn  $q$  sinkt, dann verschiebt sich gemäß Abbildung 5 die Reaktionskurve von Xavier nach innen und wird flacher, schließlich irgendwann so flach, dass sie die Reaktionskurve von Yvonne nicht (mehr) berührt. Die Effizienzwirkung von  $q$  ist beim Entscheidungsproblem von Xavier entgegengesetzt

<sup>17</sup> Man setzt in die Zielfunktionen für  $D_E$  allgemeiner  $qD_E$  ein und leitet nach  $y$  bzw.  $x$  ab.

zu dem bei Yvonne:  $q = 0$  führt zum worst-case mit  $x^* = 0$  und  $y^* = y_p$  (Inaktivität),  $q = 1$  hingegen zur sozial optimalen Entscheidung von Xavier. Xavier trägt also zum Projekt des größeren Restaurants nur insoweit bei, wie die Jurisdiktion perfekt ist.  $q$  ist in diesem Sinne ein Maß für das Durchsetzungspotential („enforcement“) des Vertrages. Vom wohlfahrtstheoretischen Standpunkt aus lässt sich folglich eine „perfekte“ Jurisdiktion (des zivilen Rechts) nicht implementieren: Damit Yvonne auf Grundlage derjenigen Reaktionskurve entscheidet, die zum sozialen Optimum (first-best) führte, muss gelten:  $q = 0$ ; Xavier hingegen wählt seinen optimalen Beitrag nur für den komplett entgegengesetzten Fall  $q = 1$  entlang der Reaktionskurve, die das soziale Optimum ermöglicht. Verbindet man die in Abbildung 5 eingetragenen Gleichgewichte für  $q = 1$ , ein hohes  $q_H < 1$ , ein niedriges  $q_L < 1$  und  $q = 0$  (Inaktivität), dann erkennt man, dass die (konvexe) Verbindungslinie die first-best Lösung  $(x^*, y_s^*)$  nicht berührt. Ist es aber vielleicht möglich, dass Yvonne und Xavier mit einem imperfekten Gericht im Hintergrund einen höheren sozialen Überschuss erzielen, als im Fall des perfekten Gerichts? Dies müsste so lange der Fall sein, bis (ausgehend von  $q = 1$ ) bei fallendem  $q$  der Wohlfahrtsverlust aufgrund eines sinkenden Anstrengungsniveaus von Xavier ( $x$  sinkt auf ein ineffizientes Niveau) durch den Wohlfahrtsgewinn aufgrund der Moral-Hazard-Reduktion bei Yvonne ( $y$  sinkt in Richtung des effizienten Niveaus) überkompensiert wird, wenn also gilt:

$$[1 - p(x)][1 - R'_p(y)] > p'(x)[R_o(y) - R_p(y)] \quad \text{bzw.} \quad \frac{[1 - R'_p(y)]}{[R_o(y) - R_p(y)]} > \frac{[p'(x)]}{[1 - p(x)]}.$$



### Abbildung 5

Die obige Ungleichheit stellt die Bedingung dar für eine „bessere“ zweitbeste Lösung.<sup>18</sup> Diese Bedingung ist umso eher erfüllt, je größer der erzielbare

Auslastungseffekt (Yvones Beitrag zum sozialen Überschuss im Zähler auf der linken Seite), je kleiner der Kapazitätseffekt (Xaviers Beitrag zum sozialen Überschuss im Nenner auf der linken Seite), je geringer der Wirkungsgrad der letzten Anstrengungseinheit von Xavier (Zähler auf der rechten Seite) ausfällt und je größer die Wahrscheinlichkeit für Vertragserfüllung bereits ist (Nenner auf der rechten Seite). Bei einer – der Tendenz nach – solchen Konstellation, bei der Xaviers potentieller Beitrag sowie sein technologischer Spielraum, einen möglichst hohen zusätzlichen Beitrag zu leisten, geringer ausfallen als die Bedeutung von Yvones Auslastungsentscheidung, ist eine juristische Institution zu teuer: Der Bedarf an Vertragsdurchsetzung ist relativ gering, die Gefahr von Fehlanreizen aufgrund des moralischen Risikos hingegen vergleichsweise groß.

## 7. Trust Responsiveness

Wie in Kapitel 2 ausgeführt, versteht man unter Trust Responsiveness (TRN) die gefühlte Verpflichtung Vertrauen zu erfüllen, weil man merkt, dass es von jemanden in einen gesetzt wurde. Eine solche Disposition weist zwei zentrale Eigenschaften auf: Erfüllt der Vertrauensnehmer nicht die in ihn gesetzten Erwartungen, entstehen ihm selbst moralische Kosten<sup>19</sup> – jedoch nur insoweit, wie er in ihn gesetztes Vertrauen registriert hat. Zur Vereinfachung und zur besseren Vergleichbarkeit mit der Wirkung formeller Institutionen nehmen wir an, diese Kosten entsprechen der Höhe des Schadensersatzes bei Nichterfüllung,  $D_E$ . Trust Responsiveness ist dann mit folgendem Kostenterm in Xaviers Zielfunktion beschreibbar:

$$t\phi D_E \quad \text{mit} \quad t \in [0, 1] \quad \text{und} \quad \phi = \min \left\{ \frac{y - y_p}{y_0 - y_p}, 1 \right\}.$$

Dabei ist  $t$  der TRN-Parameter, der angibt, wie stark diese Verhaltensdisposition bei einem repräsentativen Akteur ausgeprägt ist. In unserer Analyse ist  $t$  exogen, denn das Merkmal ist wenigstens kurzfristig nicht beeinflussbar,

<sup>18</sup> Man beachte, dass diese Bedingung nur innerhalb einer engen Umgebung unterhalb von  $(x_{DE}, y_{DE})$  – der zweitbesten Lösung bei perfekter Jurisdiktion – erfüllt sein kann. Die optimale Jurisdiktion darf in diesem Sinne nur „geringfügig imperfekt“ sein.

<sup>19</sup> Es spricht nichts dagegen, Trust Responsiveness anstatt als negativen auch als positiven Altruismus zu interpretieren. In diesem Fall wäre der Ertrag des Vertrauensnehmers das moralische Wohlempfinden, wenn er an den Vertrauensgeber das (stolze) Signal aussenden kann, dass das in ihn gesetzte Vertrauen gerechtfertigt war. Da TRN jedoch eine Verpflichtung darstellt und es – gerade in den evolutionstheoretischen Ansätzen der ökonomischen Vertrauensforschung – durchaus üblich ist, moralische Verpflichtungen in Form eines „Gewissens“ zu modellieren, bleiben wir bei der negativen Altruismus-Variante und interpretieren die moralischen Kosten der Nicht-Erfüllung als „Gewissensbisse“ (vgl. Dufwenberg / Güth, 2004).

da es z. B. kultureller Prägung unterliegt.  $\phi$  ist die TRN-Signalfunktion, die zum Ausdruck bringt, dass bei Trust Responsiveness die eigene Nichterfüllung eines gegebenen Versprechens erst dann zu einer moralischen Belastung wird, wenn zuvor registriert wurde, dass der Partner in einen Vertrauen gesetzt hat. So werden Xaviers Gewissensbisse erst dann in vollem Umfang (d. h. zu 100 %) „aktiviert“, wenn Yvonne zuvor auch in vollem Umfang Vertrauen schenkte, also  $y = y_0$  wählte. Vertraut sie Xavier hingegen nicht und beschränkt sich auf das kleinere Restaurant, dann wählt sie nur  $y = y_p$  und Xavier hat keinen Anlass, sich für die ihr dadurch entgangenen Einnahmefälle verantwortlich zu fühlen. Selbst wenn Xavier ein besonders starkes Gewissen hätte ( $t = 1$ ), dann bliebe es in diesem Fall ohne Wirkung. Der Vertrauensvorschuss des Vertrauensgebers hat im Kontext von TRN also eine wichtige Signalfunktion.<sup>20</sup>

Bevor für das TRN-Szenario die optimale Lösung bestimmt werden kann, ist zunächst eine methodische Differenzierung notwendig. Alle bisher in den Zielfunktionen von Yvonne und Xavier betrachteten Auszahlungen waren monetär und in diesem Sinn „objektiv“.<sup>21</sup> Die Erweiterung der Analyse in diesem Kapitel um altruistische Präferenzen bedeutet jedoch, dass wir es auch mit Auszahlungen zu tun haben, denen die monetäre Entsprechung fehlt. Wir behandeln dieses Problem auf die Weise, wie es im indirekt-evolutionären Ansatz von Güth/Yaari vorgeschlagen wurde und seitdem in der evolutorischen Ökonomie etabliert ist (vgl. Güth/Yaari, 1992, 23 ff.): Die ‚rein subjektiven‘ bzw. altruistischen Größen der Auszahlungsfunktion beeinflussen einerseits die Auswahl der optimalen Strategie und sind daher entscheidungsrelevant. Andererseits haben sie keine Auswirkung auf den objektiven Erfolg eines Akteurs, also auf das monetäre und damit objektiv bestimmbare Ergebnis der Interaktion.<sup>22</sup> Für die konkrete Vorgehensweise bedeutet dies, dass die altruistischen Größen vollkommen gleichwertig in die Bestimmung der Reaktionsfunktionen und damit der wechselseitig optimalen Strategiewahl eingehen, bei der Ermittlung der (diesem Strategietupel entsprechenden) realisierten Auszahlungen jedoch keine Rolle spielen.

Xaviers Zielfunktion samt Bedingung erster Ordnung lautet damit:

$$E(\pi_x) = P - c(x) - [1 - p(x)]t\phi D_E = P - c(x) - [1 - p(x)]t\phi [R_o(y) - R_p(y)]$$

$$E'(\pi_x) = p'(x)t\phi [R_o(y) - R_p(y)] = c'(x) .$$

<sup>20</sup> Man beachte den Unterschied zum Schadensersatz bei Nichterfüllung aus Kapitel 5: Vertraut Yvonne nicht, so hat Xavier bei Nichtleistung ihr dennoch den Schaden zu erstatten, der durch den ausbleibenden Kapazitätseffekt entsteht.

<sup>21</sup> Dahinter steckt die vereinfachende Annahme, dass der jeweilige Nutzen von Yvonne bzw. Xavier diesen objektiven Auszahlungen entspricht. Aus individueller und subjektiver Sicht interessieren sich also beide nur für das monetäre Ergebnis.

<sup>22</sup> „Intrinsic motivation (here guilt) can thus influence success only indirectly by its effects on choice behaviour.“ (Dufwenberg/Güth, 2004, 5).

Vergleicht man Xaviers Bedingung erster Ordnung im TRN-Szenario mit derjenigen bei perfekter Jurisdiktion aus Kapitel 5, dann erkennt man einen gemeinsamen Optimalpunkt: Bei perfektem Gewissen  $t = 1$  und vollkommener Vorleistung  $\phi(y_o) = 1$  schneidet seine Reaktionskurve diejenige von Yvonne im bekannten Punkt ( $x^* = x_{DE}, y^* = y_o$ ). Mit sinkendem  $t$  verschiebt sich seine Reaktionskurve parallel weiter nach rechts unten. Je schwächer Xaviers Gewissen ausgeprägt ist, desto höher muss Yvones Vorleistung ausfallen, damit er bereit ist, die erste Anstrengungseinheit zu leisten. Die Signalfunktion  $\phi$  verursacht gleichzeitig eine Linearisierung von Xaviers Reaktionsfunktion: Je höher Yvones Vorleistung, desto stärker das Anstrengungsniveau von Xavier, d. h. durch Trust Responsiveness wird die Beziehung zwischen Vorleistung und Anstrengung annähernd proportional.

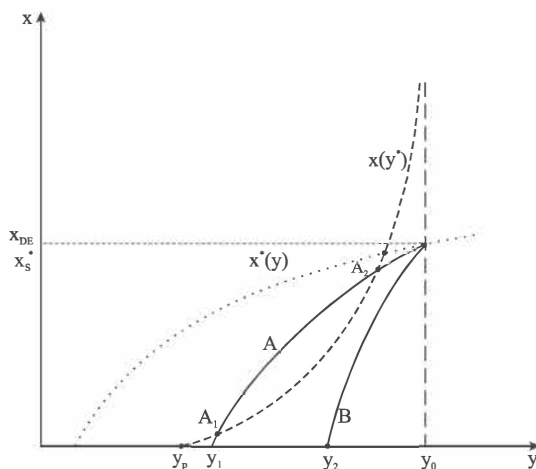


Abbildung 6

Ermöglicht nun TRN effiziente Kooperation zwischen Yvonne und Xavier? Diese Frage lässt sich nicht pauschal beantworten, sondern hängt von der Entfernung des  $y$ -Achsenabschnitts beider Reaktionskurven voneinander ab. Es ist unmittelbar einsichtig, dass im TRN-Szenario die Reaktionskurve von Xavier rechts von Yvones Kurve beginnen muss und nie links von ihr beginnen kann, da dies bedeuten würde, dass Xavier sich anstrengt ( $x > 0$ ) obwohl Yvonne noch nicht einmal in Höhe der pessimistischen Vorleistung  $y_p$  vertraut. Zunächst muß der Vertrauensnehmer ein hinreichend vertrauensvolles Signal empfangen, damit es sich für ihn lohnt, das Vertrauen zu erfüllen. Ist der  $y$ -Achsenabschnitt  $\underline{y}$  von Xaviers Reaktionskurve bzw. der Abstand der Achsenabschnitte ( $\underline{y} - y_p$ ) niedrig, dann kann es – wie aus Abbildung 6 am Achsenabschnitt  $\underline{y}_1$  erkennbar – zum Schnittpunkt beider Reaktionskurven



kommen. Der Abstand  $y - y_p$  fällt dann besonders niedrig aus, wenn nur eine geringfügige Erhöhung von Yvones Vorleistung erforderlich ist, um die erste Anstrengungseinheit von Xavier produktiv einzusetzen. Aus formaler Sicht ist dies um so eher erfüllt, je größer der folgende Term  $T$  wird:<sup>23</sup>

$$T = \frac{w'(0)}{c'_x(0)} [R'_o(y) - R'_p(y)] .$$

Dazu müssen der TRN-Parameter, der Wirkungsgrad der ersten Anstrengungseinheit sowie der Auslastungseffekt möglichst hoch ausfallen, die Grenzkosten von Xaviers erster Anstrengungseinheit  $c'_x(0)$  hingegen möglichst niedrig sein.<sup>24</sup> Kurve A in Abbildung 6 veranschaulicht diese Situation, in der aufgrund des Verlaufs von Xaviers Reaktionsfunktion multiple Gleichgewichte entstehen können. Dieses Ergebnis ist nicht überraschend, sondern veranschaulicht auf sehr originelle Weise das Merkmal zirkulärer bzw. sich selbsterfüllender Erwartungen, das bei Vertrauen eine große Rolle spielt: Yvonne versichert Xavier, ihm absolut zu vertrauen, gegeben dass er sich anstrengt. Xavier wiederum ist bereit, sich sehr anzustrengen, jedoch nur unter der Bedingung, dass Yvonne ihm absolut vertraut. Wie oft bei multiplen Gleichgewichten lässt sich die Lösung nicht immer endogen bestimmen, sondern nur exogen über Focuspunkt-Eigenschaften. Über diesen Weg ließe sich sogar ein wichtiges Konzept gesellschaftlichen Vertrauens, das Sozialkapital, in die Modellüberlegungen integrieren: In einer gesellschaftlichen Gruppe, in der eine grundsätzlich hohe Bereitschaft besteht, dem anderen mit Sympathie zu begegnen und ihm gute Absichten zu unterstellen, wäre der Punkt A<sub>2</sub> in Abbildung 6 eher erreichbar, als in einer Alternativgruppe, die ein pessimistisches und misstrauisches Weltbild hat. Diese Gruppe kann selbst bei grundsätzlich hoher Disposition zu TRN und einer potentiell sehr produktiven Beziehung in einem ‚bad equilibrium‘, wie in Punkt A<sub>1</sub> aus Abbildung 6 stecken bleiben. Fällt der Abstand der Achsenabschnitte hingegen hoch aus (in der Abbildung entspricht dies dem Abstand  $y_2 - y_p$ ), dann kommen solche Schnittpunkte nicht zustande. Die Verhaltensdisposition von Trust Responsiveness allein stellt also die Kooperation von Xavier und Yvonne noch keineswegs sicher. Dies lässt sich mit Hilfe komparativ-statischer Überlegungen am Beispiel der Kurve B in Abbildung 6 veranschaulichen: Angenommen, Xavier verspricht Yvonne, sich maximal anzustrengen ( $x_{DE}$ ). Dies veranlasst Yvonne zu einer optimalen Reaktion in Höhe von  $y(x_{DE}) < y_o$ . Eine Vorleistung in dieser Höhe aber interpretiert ein auf TRN geeichter Xavier als nicht besonders vertrauensvoll und reduziert seinen Einsatz entsprechend. Dies antizipierend passt wiederum Yvonne ihre Vorleistung nach unten an. Das Ergebnis ist Inaktivität mit  $y^* = y_p$ ,  $x^* = 0$  und

<sup>23</sup> Die Herleitung dieses Terms findet sich im Anhang.

<sup>24</sup> Bislang galt stets vereinfachend  $c'(x) = c'(y) = 1$ .

einem Wert der Beziehung von 0. Da beide annahmegemäß ihre Betriebe mit abnehmenden Skalenerträgen bewirtschaften, ist eine Mindestproduktivität Voraussetzung dafür, dass aus altruistischer Motivation heraus effiziente Kooperation entstehen kann.

## 8. Zusammenwirken von imperfekter Jurisdiktion und Trust Responsiveness

Mit Hilfe der bisherigen Ergebnisse lässt sich nun auch die Frage nach dem Zusammenwirken von formeller Institution und Vertrauen beantworten. Hierzu unterstellen wir – wie bisher – die TRN-Disposition bei Xavier und im Hintergrund eine Jurisdiktion, die bei Nichtleistung von Xavier mit Wahrscheinlichkeit  $q$  Schadensersatzzahlungen in Höhe von  $D_E$  durchsetzen kann. Xaviers Zielfunktion resp. Bedingung erster Ordnung lauten für dieses Szenario:

$$E(\pi_x) = P - c(x) - [1 - p(x)](t\phi + q)D_E$$

$$E'(\pi_x) = p'(x) \left[ \underbrace{t\phi D_E}_{(a)} + \underbrace{qD_E}_{(b)} \right] = c'_x(x) .$$

Wir beschränken uns für die Interpretation auf den interessanten Fall der Kurve B aus Abbildung 6, bei dem selbst ein perfekt ausgeprägtes Gewissen bei Xavier wirkungslos blieb.

Der Grund dafür lag in der Signalfunktion  $\phi$  in Term (a) der oben abgebildeten Bedingung erster Ordnung: Das Gewissen aktiviert sich nur bei einer Vorleistung von Yvonne  $y > y_p$ , aber erst ab einer Vorleistung von  $y_2$  sind Auslastungseffekt und Gewissen für Xavier hoch genug, damit er sich überhaupt bewegt und die erste Anstrengungseinheit einsetzt. Bei einer Vorleistung  $y_2$  erwartet Yvonne jedoch einen deutlich höheren Einsatz und bleibt damit lieber beim kleineren Restaurant. Mit einem Gericht im Hintergrund kommt bei Xaviers Entscheidung jedoch auch Term (b) zum Tragen: Wenn Xavier sein Versprechen nicht hält, dann hat er mit Wahrscheinlichkeit  $q$  den Kapazitätseffekt als Schadensersatz zu entrichten. Das bedeutet aber, dass für ihn – wie aus Kapitel 5 bekannt – auch dann Schadensersatzzahlungen anfallen, wenn Yvonne ihm nicht vertraute und sich stattdessen mit  $y = y_p$  auf das kleinere Restaurant einstellte. Potentielle Schadensersatzzahlungen machen für Xavier die erste Anstrengungseinheit früher attraktiv (d. h. Nichtanstrengung ist im Verhältnis dazu nun teurer geworden) und seine Reaktionskurve verschiebt sich mit steigendem  $q$  nach oben (vgl. Abbildung 7).

Die Wirkung des Gerichts auf Yvannes Entscheidung ist auf Grund des Versicherungseffektes entgegengerichtet und wurde bereits in Kapitel 5 formal und grafisch diskutiert. Man erkennt an der Grafik, dass eine Schadensersatz-

regel schon bei niedrigem Perfektionsgrad  $q$  der Jurisdiktion das Inaktivitätsproblem überwindet: Da Xavier auch schon dann den Anreiz hat, sich anzustrengen, wenn Yvonne noch nicht vertraute, beginnt seine Reaktionskurve früher als die von Yvonne. Für den durchgehend unterstellten Verlauf der Funktionen bedeutet dies, dass effiziente Kooperation möglich ist (s. Punkt A in Abbildung 7). Interessanterweise kann eine zu perfekte Institution (z. B.  $q = q_H$ ) sogar kontraproduktiv wirken, da Xavier Anreize zur „Übererfüllung“ bekommt (Punkt B in Abbildung 7). Ist das Gewissen von Akteuren verhältnismäßig stark ausgeprägt, dann ist eine nahezu perfekte Institution zur Vertragsdurchsetzung nicht nur nicht erforderlich, sondern sogar kontraproduktiv, weil ökonomisch ineffizient. Ab einer gewissen Mindestausgestaltung der formellen Institution besteht zwischen Moral (hier als Vertrauen im engeren Sinn) und formellen Institution im Hintergrund (als Vertrauen im weiteren Sinn) der Tendenz nach eine substitutive Beziehung. Eine Mindestausgestaltung der Institution ist – zumindest im TRN-Kontext – jedoch erforderlich: Die Institution übernimmt für die Kooperation von Xavier und Yvonne eine Katalysatorfunktion, indem sie bei Xavier für den notwendigen Anstoß sorgt, eine gewisse „Anschubleistung“ zu erbringen, zu der er im Szenario mit perfektem Gewissen aber ohne Institution keinen Anreiz hätte. Bereits eine Institution mit geringem  $q$  kann hierfür schon ausreichend sein. In dieser Hinsicht besteht zwischen Vertrauen und formellen Institutionen ein komplementärer Zusammenhang.

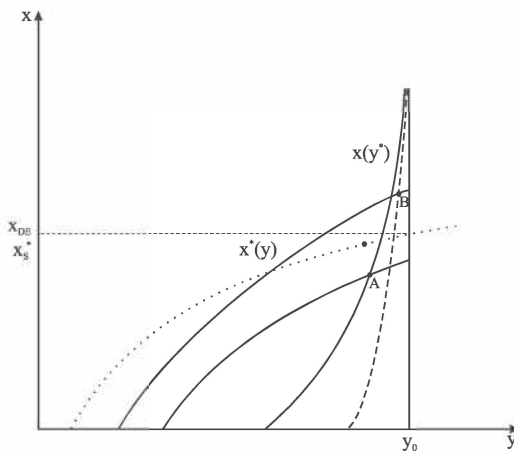


Abbildung 7

## 9. Fazit

Die eingangs aufgeworfene und in der einschlägigen Literatur kontrovers diskutierte Frage – ob Vertrauen und formelle Institutionen (insb. „gerichts-feste“ Verträge mit Sanktionsklauseln) sich zueinander eher als Substitut oder als Komplement verhalten – lässt sich mit Rückgriff auf die Theorie des Schadensersatzes und dem Konzept der „trust responsiveness“ durchaus beantworten.

Die „traditionelle“ Vertrauensforschung hat gezeigt, dass es in einem einmaligen Spiel („trust game“) – ohne jegliche Absicherung oder sonstige weitergehende Annahmen (etwa Reputationseffekte, Triggerstrategien etc.) nicht zu der denkbaren erstbesten Lösung kommen wird, die den sozialen Überschuss maximiert. In einer solchen Situation gibt es keinerlei Sicherheit (für keine der Parteien), dass gegebenes Vertrauen vom Gegenüber honoriert werden wird. Im Gegenteil: Wenn sicher antizipiert werden kann, dass gegebenes Vertrauen enttäuscht werden wird, dann wird es gar nicht erst geschenkt. Dieses Dilemma lässt sich auch nicht durch einen anreizkompatiblen Vertrag lösen. Insoweit ist ein solcher Vertrag auch kein Substitut zu Vertrauen. Von den drei gebräuchlichsten Schadensersatzklauseln ist (nur) die Regel, welche für die Zahlung des individuell geregelten Schadensersatzes Sorge trägt, in der Lage, die Parteien zur Realisierung der sozial optimalen Lösung (optimales Vertrauensniveau vs. optimales Anstrengungsniveau) zu bewegen. Die Bestimmung der „optimalen Schadenshöhe“ setzt allerdings gerade die Kenntnis der von beiden Parteien im Konsens festgelegten erwarteten Erträge aus dem gemeinsamen Projekt voraus und verursacht – ganz im Sinne von R. Coase – Transaktionskosten. Eine sichere Erwartung über zu leistenden bzw. zu empfangenden Schadensersatz setzt perfekte Gerichte bzw. perfekte Jurisdiktion voraus. Die Konstanz und sogar eine Erhöhung des sozialen Überschusses kann allerdings auch bei reduziertem Perfektionsgrad des Gerichts möglich sein, nämlich dann, wenn der Vertrauensgeber den größeren Teil des sozialen Überschusses verantwortet. Dieser Sachverhalt verdeutlicht gleichzeitig, dass Vertrauen stets als zweiseitiges Phänomen verstanden werden muss. Ein signifikanter Defekt der Jurisdiktion führt auf jeden Fall zu einem Wohlfahrtsverlust für beide Parteien. Sie werden dann gezwungen sein, auf Mechanismen privater Vertragsdurchsetzung zurückzugreifen, welche hohe Transaktionskosten verursachen und z. T. stark von Misstrauen geprägt sind (Screening, Stigmatisierung etc.).

Wird allerdings der Anreizmechanismus der Beteiligten um den Aspekt der „trust responsiveness“ (TRN) erweitert, so findet man durchaus andere Ergebnisse: Effiziente Kooperation kann sich einstellen, sobald der Vertrauensnehmer zu TRN veranlagt ist, er ein entsprechendes Signal des Vertrauensgebers empfängt und das gemeinsame Projekt eine Mindestproduktivität aufweist. Dabei können multiple Gleichgewichte auftreten, von denen aus Sicht der

Maximierung des sozialen Überschusses nicht alle wünschenswert sind. Formelle Institutionen können in diesem Zusammenhang folgende Rolle übernehmen: Sie sind insofern komplementär als dass sie – in mäßiger Intensität – im Zusammenspiel mit TRN den Vertrauensnehmer eher zur Kooperation bzw. Vertragserfüllung animieren als wenn es sie nicht gäbe. TRN und Gerichte zusammen müssen dabei nur in einem Mindestmaß ausgeprägt sein. Ist diese Voraussetzung erfüllt, dann ist zusätzlich erreichbare Kooperation nur effizient, wenn eines von beiden verstärkt wird, d. h. es besteht ein eher substitutiver Zusammenhang zwischen der „intrinsischen“ Bereitschaft zur Kooperation einerseits und den Anreizen zur Vertragserfüllung, die von einem Gericht ausgehen, andererseits.

### Anhang: Herleitung der Bedingung für den Abstand der Y-Achsenabschnitte beider Reaktionskurven aus Kapitel 7

Der Achsenabschnitt von Yvones Reaktionskurve ist mit  $y_p$  festgelegt. Eine Vorleistung in Höhe von  $y_p$  maximiert den Gewinn im kleinen Restaurant, d. h.  $y_p$  löst die Bedingung

$$(1) \quad R'_p(y) = c'(y) .$$

Der Y-Achsenabschnitt von Xaviers Reaktionskurve  $\underline{y}$  (= Mindeststärke des Vertrauenssignals, bei der die erste Anstrengungseinheit optimal zum Einsatz kommt) ist per Definition  $x^*(\underline{y}) = 0$ , d. h.  $\underline{y}$  löst Xaviers Bedingung 1. Ordnung für  $x = 0$ :

$$(2) \quad p'(0)t\phi[R_o(y) - R_p(y)] = c'(0) .$$

Im Punkt  $y_p$  wird für Xavier das Verhältnis von Grenzerlös und Grenzkosten Null (wegen  $\phi(y_p) = 0$ ), d. h. bei derjenigen Vorleistung, bei der der Gewinn im kleinen Restaurant maximiert wird, ist Xaviers Grenzgewinn strikt negativ:

$$(3) \quad \phi(y_p)[R_o(y_p) - R_p(y_p)] \frac{tp'(0)}{c'(0)} = 0 < \frac{R'_p(y_p)}{c'(y_p)} = 1 .$$

Mit sukzessiv steigendem  $\underline{y}$  wird sein Grenzgewinn Null und es lohnt sich für ihn, sich Mühe zu geben:

$$(4) \quad \phi(\underline{y})[R_o(\underline{y}) - R_p(\underline{y})] \frac{tp'(\underline{y})}{c'(\underline{y})} = 1 > \frac{R'_p(\underline{y})}{c'(\underline{y})} .$$

Der Abstand zwischen beiden Reaktionskurven ist um so kleiner, je geringer man  $\underline{y}$  variieren muss, um von Situation (3) zur Situation (4) zu gelangen. Die Ableitung der linken Seite von Gleichung (4) nach  $\underline{y}$  liefert den Term  $T$  aus Kapitel 7.

## Literatur

- Albert, M.* (1999): Kommentar zu T. Ripperger, Die Effizienz des Vertrauensmechanismus bei der Organisation internationaler Transaktionen, in: K.-E. Schenk et al. (Hrsg.), *Globalisierung und Weltwirtschaft*, Tübingen, 292–301.
- Berg, J. et al.* (1995): Trust, Reciprocity, and Social History, *Games and Economic Behaviour* 10, 122–142.
- Cochard, F./Nguyen Van, P./Willinger, M.* (2004): Trusting Behavior in a repeated investment game, *Journal of Economic Behavior & Organization* 55, 31–44.
- Cooter, R./Ulen, T.* (1988): *Law and Economics*, Scott, Foreman and Company, Illinois / London (England).
- Dasgupta, P.* (1988): Trust as a Commodity, in: D. Gambetta (ed.), *Trust. Making and Breaking Cooperative Relations*, New York, 49–72.
- Demougin, D.* (1999): Kommentar zu T. Ripperger, Die Effizienz des Vertrauensmechanismus bei der Organisation internationaler Transaktionen, in: K.-E. Schenk et al. (Hrsg.), *Globalisierung und Weltwirtschaft*, Tübingen, 302–310.
- Dufwenberg, M.* (2002): Marital investments, time consistency and emotions, *Journal of Economic Behavior & Organization* 48, 57–69.
- Dufwenberg, M./Güth, W.* (2004): The Psychological Game of Trust, *Papers on Strategic Interaction*, Max Planck Institute for Research into Economic Systems 19.
- Fehr, E./List, J.* (2004): The Hidden Costs and Returns of Incentives – Trust and Trustworthiness among CEOs, *Journal of the European Economic Association* 2 (5), 743–771.
- Guerra, D./Zizzo, D. J.* (2004), Trust responsiveness and beliefs, *Journal of Economic Behavior & Organization* 55, 25–30.
- Güth, W./Yaari, M.* (1992): An Evolutionary Approach to Explaining Reciprocal Behavior in a Simple Strategic Game, in: U. Witt (ed.), *Explaining Process and Change, Approaches to Evolutionary Economics*.
- Guiso et al.* (2000): *The Role of Social Capital in Financial Development*. NBER Working Paper No. 7563, Cambridge (Massachusetts).
- Lorenz, E.* (1999): Trust, Contract and Economic Cooperation, *Cambridge Journal of Economics* 23, 301–315.
- Milgrom, P./Roberts, J.* (1992): *Economics, Organisation and Management*, Englewood Cliffs / New York.
- Nooteboom, B.* (2002): *Trust. Forms, Foundations, Functions, Failures and Figures*, Cheltenham (UK) / Northampton (MA, USA).
- Richter, R./Furubotn, E. G.* (1999): *Neue Institutionenökonomik*, 2. Auflage, Tübingen.
- Rogerson, W. P.* (1984): Efficient Reliance and Damage Measures for Breach of Contract, *The Rand Journal of Economics* 15, 39–53.
- Schweizer, U.* (2003): *Cooperative Investment Induced by Contract Law*, German Working Papers in Law and Economics 2003, Paper 25, University of Bonn, Germany.

- Sell, F. L.* (2005): Vertrauen und Vertrauenserosion: Ökonomische Funktionen und Effekte, in: M. Held/G. Kubon-Gilke/R. Storn (Hrsg.), *Reputation und Vertrauen*, Jahrbuch Normative und institutionelle Grundfragen der Ökonomik, Bd. 4, Marburg, 161 – 184.
- Sell, F. L.* (2004): Vertrauen: Auch eine ökonomische Kategorie, in: G. Blümle et al. (Hrsg.), *Perspektiven einer kulturellen Ökonomik*. Kulturelle Ökonomik, Band 1, Münster/Hamburg/Berlin/London, 399 – 410.
- Stiglitz, J. E.* (2002): Die Schatten der Globalisierung, German Edition of „The Shadows of Globalisation“, Berlin.
- Zak, P. J./Knack, S.* (2001): Trust and Growth, *The Economic Journal* 111, April, 295 – 321.