2004

Market Design with Endogenous Preferences

Aviad Heifetz
aviadhe@openu.ac.il

Ella Segev
Ben-Gurion University of the Negev, ellasgv@bgu.ac.il

Eric L. Talley
Columbia Law School, etalley@law.columbia.edu

CLEO RESEARCH PAPER SERIES LAW & ECONOMICS RESEARCH PAPER SERIES

Market Design with Endogenous Preferences

Aviad Heifetz[*]   Ella Segev[†]   Eric Talley[‡]

April 2004

Abstract

This paper explores the interdependence between market structure and an important class of extra-rational cognitive biases. Starting with a familiar bilateral monopoly framework, we characterize the endogenous emergence of preference distortions during bargaining which cause negotiators to perceive their private valuations differently than they would outside the adversarial negotiation context. Using this model, we then demonstrate how a number of external interventions in the structure and/or organization of market interactions (occurring before trade, after trade, or during negotiations themselves) can profoundly alter the nature of these dispositions. Our results demonstrate that many such interventions frequently (though not always) share qualitatively similar characteristics to market interventions that are often proposed for overcoming more conventional forms of market failure. Nevertheless, our analysis underscores the importance of understanding the precise link between cognitive failures and market structure prior to the implementation any particular proposed reform.

Keywords: Market Design, Endogenous Preferences
JEL classification numbers: H20, D40, Z13

1   Introduction

In recent years, the rational actor model — pivotal to much of modern economic theory — has fallen under renewed criticism from scholars both inside and outside economics proper. By at least some accounts, this scrutiny is long overdue: Indeed, there are scores of examples where observed behavior appears strikingly at odds with at least the most straightforward predictions of models with strong rationality assumptions. The growing literature in *behavioral economics* is largely dedicated to cataloguing and systematizing instances in which preferences are internally inconsistent, dynamically unstable, or actuarially biased. Examples of such phenomena include overconfidence, endowment effects, framing effects, self-serving biases, heuristics, cycling, and various forms of bounded rationality (see e.g. Rabin (1998) for an overview).

Given the ascendancy of this literature, there is a natural urge to transcend the positive connections between decision-making problems and cognitive biases, and to explore the *normative* consequences that such phenomena imply. Not surprisingly, a number of recent efforts in the literature appear to do just that, using topics within behavioral economics as springboards for proposing market interventions or legal reforms that attempt to compensate for the existence of cognitive preference distortions. Sunstein (2002) for example, considers how elimination of at-will employment doctrine may help address problems with endowment effects. [1]

These accounts view cognitive biases as *an exogenous assumption* within a behavioral model, and take preferences (distorted by biases) to be a primitive building block of equilibrium behavior. Such an assumption, however, stands in contrast with the experimental evidence, which suggests that many cognitive dispositions appear to be highly context specific, rising to first-order importance in certain settings, while curiously marginal in others (Camerer et al. 2003). Thus, without a more general theory of context, it is difficult to predict how (or whether) various biases occurring in the laboratory should translate to the real world targets of policy reforms, and how these biases may be *affected* by such interventions.

In this paper, we propose a model for analyzing how context and cognition plausibly interact with one another, and a resulting framework for studying institutional design within such a setting. Our analysis reveals that the task of designing institutions in the presence of cognitive biases is somewhat more complicated than in the classical approach to design problems, for at least two reasons. First (and most centrally), regulatory interventions *themselves* are likely to distort context, and in so doing may affect the direction or magnitude of various cognitive dispositions. When contemplating issues of institutional design and policy, then, one must take care not only to identify the biases which cause inefficiencies, but also to anticipate the *feedback effects* induced by the very

---

[1]Other examples include Jolls (1998) who argues that wealth redistribution is more efficiently accomplished through legal liability rules than through tax and trasnfer systems, since people systematically underestimate the likelihood of legal liability.

regulatory apparatus meant to compensate for them. Such feedback effects are frequently not incorporated into normative policy proposals, and their omission could very well lead to imprecise, inefficient, and ultimately ineffectual reforms.

Second, even if one could anticipate the feedback effects described above, a particularly thorny problem remains in specifying a reasonable definition of "optimality" in the presence of endogenous preferences. Indeed, conventional notions of economic welfare become more contested in environments where preferences themselves shift over time (see, e.g., Carmichael & MacLeod (2002)). As we demonstrate below, the nature of an optimal regulatory intervention may turn dramatically on whether one defines optimality in terms of a context-specific or context-independent definition of preferences.

The analytic model we propose is one in which cognitive dispositions — much like behavior itself — arise endogenously, through an equilibrium process. This approach draws from a recent literature[2] whose origins trace back (at least) to Becker (1976), and which hinges on the *commitment value* of preferences: i.e., it may "pay" to be concerned with motives other than one's own wealth, since so doing can induce other actors to make favorable accommodations in their equilibrium behavior. Ultimately, holding the economic environment constant, the interaction between individual biases and this responsive accommodation by others can generate a stable equilibrium both in preferences and in behavior. Accordingly, when the underlying economic environment governing the interaction changes, preferences and behavior will both adjust as well.

The linchpin of the equilibration process we posit is an assumption that those who adopt preference dispositions yielding larger material rewards (as measured by their context-independent preferences) also tend to become more prominent in the population of players. While this equilibrating process is certainly reminiscent of literal evolutionary equilibrium concepts, it is significantly broader than that. The same account, for example, would also apply to situations where individuals simply imitated and adopted the attitudes and norms of those who appear to be successful over the long term, thereby reducing the economic influence of others. This concept of preference equilibrium is a natural embarking point for an economic analysis of preference endogeneity, since (1) it is grounded in first principles rather than exogenous assumptions about biases; and (2) it ultimately subscribes to the notion that individuals adapt in a way which is beneficial to their own, genuine welfare (albeit indirectly and unconsciously).

In order to illustrate the application of our framework, we explore its consequences in what is perhaps the most fundamental arena of economic interaction: bilateral exchange. Using a familiar, canonical framework of noncooperative bargaining with two-sided pri-

---

[2]Samuelson (2001) presents a brief overview of this literature, while Dekel, Ely and Yilankaya (1998) and Heifetz, Shannon and Spiegel (2003) present general results in this vein. Examples include Güth and Yaari (1992), Huck and Oechssler (1998), Fershtman and Weiss (1997, 1998), Rotemberg (1994), Bester and Güth (1998), Possajennikov (2000), Bolle (2000), Bergman and Bergman (2000), Koçkesen, Ok, and Sethi (2000a, 2000b), Guttmann (2000), Sethi and Somanathan (2001), Kyle and Wang (1997), Benos (1998), Heifetz and Segev (2001).

vate information as a benchmark (e.g., Myerson & Satterthwaite 1983), we characterize the emergence of preference distortions during bargaining that cause negotiators to skew their perceived private values away from those they would perceive outside the bargaining context. Such preference distortions are commonly observed in the experimental literature, often associated with the "endowment effect," the "self-serving bias," or both. We demonstrate how such cognitive dispositions can benefit private negotiators, effectively transforming them into "tougher" bargainers than they would be in the absence of bias, thereby augmenting the credibility of their threat to exit without an agreement. Moreover, based on the analysis in Heifetz and Segev (2001) we illustrate how such transitory preference distortions are a viable equilibrium trait within a population of parties who bargain in thin market settings, identifying the emerging preference-behavior equilibrium as a function of the bargaining scheme.

We then turn our attention to the question of optimal institutional design. Using the bilateral bargaining framework described above as a template, we demonstrate how various market interventions — either by the state or by a benevolent third party — can have profound effects on both the existence and magnitude of transitory preference distortions during negotiation. Accounting for these effects can cause an optimal regulatory scheme to differ (sometimes dramatically) from that in which cognitive dispositions were either assumed away or treated as exogenous primitives. To our knowledge, this is the first systematic attempt to consider the prescriptive interrelationship between cognition, context, and market design.

Significantly, the market interventions we analyze are not merely fanciful figments of our collective imagination, but rather real-world mechanisms through which third parties can (and do) exercise regulatory power at various points in the bargaining process. In particular, we focus on three genres of actual market intervention (differentiated by the time at which regulation occurs) that strike us as particularly salient:

- **Ex post intervention.** A number of institutional devices exist for rewarding traders upon reaching a negotiated outcome. For example, various elements of the tax code often act to subsidize the consummated bargains.[3] This approach is also increasingly common in the international arena, as funding sources (such as the World Bank) have begun to de-emphasize the importance of demonstrating economic need, basing their funding decisions more centrally on a model of rewarding the resolution of international conflicts and the implementation of internal agreements to distribute aid effectively.[4]

- **Interim intervention.** Other forms of regulatory intervention occur at the negotiation process itself, artificially constraining the types of bargains that are allowed.

---

[3]See, e.g., Cal. Rev. & Tax Code § 17053.84 (paying a 15% tax credit for the purchase and installation of solar energy equipment by California residents).
[4]See, e.g., Dollar (2000).

In contract law, for example, numerous legal rules (such as the doctrines of consideration and unconscionability in contract law, and the doctrine of moeities in admiralty law) operate to narrow the range of enforceable bargaining outcomes relative to what the parties would find individually rational. In addition, in some circumstances price/wage ceilings and floors operate with a similar constraining effect.

- **Ex ante intervention.** Still other regulatory interventions take place before bargaining even begins, at the point in which initial property rights are assigned. A substantial portion of common law doctrines and statutory provisions are dedicated to specifying the contours of individual property rights, ranging from strong monolithic entitlements protected by injunctive relief, to weak entitlements that are either protected solely with damages or are subject to joint ownership (e.g., Ayres & Talley 1995). Individuals frequently negotiate transfers of title in the shadow of these entitlements.

Within each of these examples, we show how an optimal regulatory intervention would account not only for garden variety market failures, but also for the endogenous cognitive shifts that regulatory interventions themselves can trigger. In so doing, we highlight how many of the now-accepted approaches for mitigating *strategic* barriers to trade might fare once *cognitive* barriers are also taken into account.

In some instances, the fit is a poor one. For example, while subsidizing successful trades (i.e., ex post intervention) has long been recognized as a method for counteracting the effects of strategic behavior by privately-informed parties, we demonstrate that such subsidy schemes can themselves aggravate transitory preference shifts. As a result, optimal market design in the presence of endogenous cognitive biases may require a significantly greater amount of intervention in the market (i.e., a higher subsidy rate) than would be necessary in an environment where agents were devoid of such distortions.[5]

Less pessimistically, we demonstrate that there are some forms of intervention in which strategic and cognitive concerns overlap. For instance, we show how "weak" property entitlements (such as joint ownership or "fuzzy" property rights) not only help to mitigate strategic misrepresentation (e.g., Cramton, Gibbons & Klemperer (1987); Ayres & Talley (1995)), but they can also help to dissipate cognitive dispositions towards toughness.

In a similar vein, we find that interim interventions constraining the types of allowed bargains, typically devised to secure incentives in the process of information exchange, also help in mitigating transitory cognitive shifts. However, it turns out that *excessive* such intervention – *beyond* the extent prescribed in Myerson and Satterthwaite (1983) –

---

[5]This particular result turns on the social planner attempting to maximize a-contextual preferences. However, as we show below, regardless of what the social planner attempts maximize, the optimal subsidy will generally diverge (either above or below) that of a classical analysis in which biases are assumed away.

will in fact make the traders happiest in expectation *given* the *endogenous* level of these shifts.

Our analysis illustrates also the second obstacle in conducting market design in the presence of cognitive biases: the elusive meaning of the term "optimal." Indeed, in a situation where individual preferences are prone to endogenous shifts, utilitarian notions such as efficiency become significantly more indeterminate than they are in traditional rational choice theory. In particular, one might justifiably choose to focus on a notion of welfare rooted in a-contextual preferences (what we shall intermittently refer to as "wealth"), corresponding to those preferences individuals manifest in the abstract, outside of an adversarial bargaining context. Alternatively, one might conceive of welfare rooted in their preferences during trade (what we shall intermittently refer to as "happiness"), corresponding to those that individuals would perceive through their transitory preference dispositions at the point of bargaining. As we demonstrate below, this distinction is an important issue for policy design, as wealth and happiness maximizing approaches frequently point at divergent institutional structures.

The rest of this paper is organized as follows. Section 2 introduces a general framework for institutional design problems in the presence of endogenous dispositions. Section 3 applies this framework to the case of bilateral exchange with private information, as in Heifetz and Segev (2001). We start by deriving the dependence of the equilibrium dispositions on the bargaining mechanism, and demonstrate how this idea is operationalized in a particular bargaining equilibrium. Section 4 then turns to our constructive enterprise, demonstrating (ad seriatim) how ex post, interim, and ex ante interventions into market structure can affect the existence and degree of preference distortions during bargaining. We characterize the optimal regulatory intervention under the alternative goals of maximizing actual gains from trade ("wealth") versus maximizing the gains from trade as the traders perceive them to be during trade (the traders' "happiness"). We compare not only these optimal interventions to one another, but also against the baseline case in which cognitive dispositions were wholly absent. Section 5 concludes, and the Appendix provides the proofs of the propositions.

2 General Framework: the design of institutions when design affects preferences

In this section we describe a general framework to evaluate institutional design when preferences of individuals may be endogenously sensitive to this design. The next two sections apply these ideas to the case of bilateral exchange with private information.

Let O be a set of outcomes pertaining to the individuals $i \in I$. Let $\mathbb{U}_i$ be a set of utility functions $U_i : O \to R$ that individual i may have. We denote by $U = (U_i)_{i \in I}$ a utility profile of the individuals. The set of utility profiles is $\mathcal{U} = \prod_{i \in I} \mathbb{U}_i$.

Let there be a welfare aggregation function $\mathcal{W} : \mathcal{U} \to \mathbb{U}$, where $\mathbb{U}$ is also a set of utility functions $U : O \to R$. That is, for each utility profile U, $\mathcal{W}(U) : O \to R$ is a utility function itself.

The strategies available to individual $i \in I$ are S_i, and $S = \prod_{i \in I} S_i$ is the set of possible strategy profiles. A mechanism is a function $\mu : S \to O$, which specifies an outcome for each strategy profile of the individuals. When the individuals have the utility profile U, a strategy profile $s^* \in S$ is a Nash Equilibrium of the mechanism μ if for each individual $i \in I$

$$U_i\left(\mu\left(s^*\right)\right) \geq U_i\left(\mu\left(s_i, s^*_{-i}\right)\right) \tag{2.1}$$

for every strategy $s_i \in S_i$. As usual, $\left(s_i, s^*_{-i}\right)$ denotes the strategy profile obtained from s^* by replacing only the strategy s^*_i of individual i by s_i.

Let $\mathcal{M}$ be the set of available mechanisms. We assume that for every utility profile $U \in \mathcal{U}$ there exists a mechanism $\mu \in \mathcal{M}$ which has a Nash equilibrium for U. For every mechanism $\mu \in \mathcal{M}$ for which this is the case, we assume that one Nash equilibrium for U is singled out, and denoted by $s^*(U, \mu)$.

An institutional design is a map $\mathfrak{D} : \mathcal{U} \to \mathcal{M}$ such that for every utility profile $U \in \mathcal{U}$, the mechanism $\mathfrak{D}(U)$ has a Nash equilibrium for U. The induced outcome is then

$$o(U, \mathfrak{D}) = \mathfrak{D}(U)\left(s^*\left(U, \mathfrak{D}(U)\right)\right) \tag{2.2}$$

We denote by $\mathcal{D}$ the collection of available designs.

For a given utility profile U, the classical problem of institutional design consists of choosing a design $\mathfrak{D} \in \mathcal{D}$ so as to maximize $\mathcal{W}(U)(o(U, \mathfrak{D}))$ for every $U \in \mathcal{U}$. In words, the challenge is to find, for every utility profile $U \in \mathcal{U}$, the mechanism such that the outcome induced by that mechanism at its Nash equilibrium will maximize the aggregate welfare.

However, given the abundance of evidence on the sensitivity of preferences to context, it may very well be the case that the design itself also has an *unconscious effect* on the individuals' preferences. We therefore assume that in the context induced by the design $\mathfrak{D}$, an individual with a utility function U_i will unconsciously try to maximize $U_i^{\mathfrak{D}}$ rather than U_i. Thus, the institution will ultimately impose the mechanism $\mathfrak{D}\left(U^{\mathfrak{D}}\right)$ rather than $\mathfrak{D}(U)$, and the implemented outcome will be $o\left(U^{\mathfrak{D}}, \mathfrak{D}\right)$ rather than $o(U, \mathfrak{D})$.

How does the utility profile adapt to the institutional design $\mathfrak{D}$? That is, how is the map $U \to U^{\mathfrak{D}}$ determined? If we are to assume an unconscious adaptation of preferences to context, it is first of all natural to assume that such an adaptation cannot be too "wild", but rather constrained to some "neighborhood" of specific distortions of the original preferences. Formally, we therefore assume that for each utility function $U_i \in \mathbb{U}_i$ there corresponds a set of utility distortions $\mathcal{N}(U_i) \subseteq \mathbb{U}_i$, which contains U_i and from which $U_i^{\mathfrak{D}}$ can emerge.

How is $U_i^{\mathfrak{D}}$ ultimately singled out of $\mathcal{N}(U)$? Here it is natural to assume that while the distortion from U_i to $U_i^{\mathfrak{D}}$ is not the result of a conscious process, $U_i^{\mathfrak{D}}$ eventually adjusts

so as to maximize the base utility U_i given the institution $\mathfrak{D}$ and the emerging utility functions $U_j^{\mathfrak{D}}$ of the other individuals $j \neq i$. Formally, denote by $\left(\tilde{U}_i, U_{-i}^{\mathfrak{D}} \right)$ the utility profile one obtains from $U_i^{\mathfrak{D}}$ by replacing only the utility function $U_i^{\mathfrak{D}}$ of individual i by $\tilde{U}_i$. Then we say that $U^{\mathfrak{D}}$ is a *preference equilibrium* utility profile within the institutional design $\mathfrak{D}$ if for every individual $i \in I$

$$U_i^{\mathfrak{D}} \in \arg \max_{\tilde{U}_i \in \mathcal{N}(U)} U_i \left(o \left(\left(\tilde{U}_i, U_{-i}^{\mathfrak{D}} \right), \mathfrak{D} \right) \right) \tag{2.3}$$

In other words, $U^{\mathfrak{D}}$ is a Nash equilibrium of a meta-game with the strategy space $\mathcal{N}(U_i)$ for individual $i \in I$ and the payoff function $f_i : \prod_{i \in I} \mathcal{N}(U_i) \to R$ defined by

$$f_i \left(\tilde{U} \right) = U_i \left(o \left(\tilde{U}, \mathfrak{D} \right) \right) \tag{2.4}$$

Typically, the preference equilibrium utility profile $U^{\mathfrak{D}}$ indeed varies with $\mathfrak{D}$ and is different than U. Intuitively, this is because the distortion from U_i towards $U_i^{\mathfrak{D}}$ shifts the Nash equilibrium behavior of the other individuals from $s_{-i}^* \left(\left(U_i, U_{-i}^{\mathfrak{D}} \right), \mathfrak{D} \left(U_i, U_{-i}^{\mathfrak{D}} \right) \right)$ to $s_{-i}^* \left(U^{\mathfrak{D}}, \mathfrak{D} \left(U^{\mathfrak{D}} \right) \right)$, both directly – through the first argument of s_{-i}^*, and also indirectly, via the effect on the mechanism implemented in the institution $\mathfrak{D}$ – the second argument of s_{-i}^*. This effect on others' equilibrium behavior may more than compensate the individual for the fact that her own equilibrium choice $s_i^* \left(U^{\mathfrak{D}}, \mathfrak{D} \left(U^{\mathfrak{D}} \right) \right)$ maximizes $U_i^{\mathfrak{D}}$ rather than her genuine U_i.

This approach to endogenous preferences is by now well established in the literature. It dates back at least to the seminal approach of Becker (1976) (e.g. the "rotten kid theorem", which exemplified the commitment value of altruism), and elaborated further in the numerous contributions cited in footnote 2. In particular, many of these contributions analyzed the evolutionary viability of equilibrium preferences, and showed that they are either evolutionary stable in the space of preferences represented by the utility functions in $(\mathcal{N}(U_i))_{i \in I}$, or, even stronger, the sole survivors in any regular payoff-monotonic selection process in which preferences with higher fitness (with the fitness function f_i in (2.4)) proliferate at the expense of less fit preferences.

This effect of the institutional design on the very preferences that the individuals try to maximize raises a new question regarding the object of social maximization. Should the design be chosen so as to maximize $\mathcal{W}(U) \left(o \left(U^{\mathfrak{D}}, \mathfrak{D} \right) \right)$ or rather $\mathcal{W} \left(U^{\mathfrak{D}} \right) \left(o \left(U^{\mathfrak{D}}, \mathfrak{D} \right) \right)$? The former approach is based on the assumption that the preference shifts from U_i to $U_i^{\mathfrak{D}}$ are transitory and short-lived, and welfare should be evaluated according to the base utility profile U. The latter approach aims at maximizing the aggregate welfare of the individuals according to their utility profile $U^{\mathfrak{D}}$, i.e. according to their preferences when they interact within the institution. The analysis in this paper will address both these approaches.

3 Bilateral Exchange with Endogenous Preferences

We now proceed to analyze how the general framework of the previous section can be applied to the case of bilateral exchange with private information, as in Heifetz and Segev (2001).

Consider a bilateral monopoly between a potential seller (denoted as S) and buyer (denoted as B), who bargain over an undifferentiated good or legal entitlement. Both parties possess private information about their true valuations of the entitlement, but it is commonly known that these core valuations are drawn independently from uniform distributions on the same support $[\underline{a}, \bar{a}]$, which we normalize to be $[0, 1]$.[6] When we say that the core valuation of the seller is s, we mean that this is the minimal price for which she would be willing to sell the good were she to trade in some market as a price-taker. Similarly, the core valuation b of the buyer is the maximal price he would be willing to pay for the good as a price-taker.

The bargaining scheme, however complicated, eventually gives rise to some probability of trade $p(s, b)$ for each pair of seller and buyer valuations $s, b \in [0, 1]$, and an average monetary transfer $t(s, b)$ from the buyer to the seller given trade. The ex ante probability of trade is therefore given by

$$P = \int_0^1 \int_0^1 p(s, b) ds db, \tag{3.1}$$

and the ex ante expected gains from trade are

$$G = \int_0^1 \int_0^1 (b - s) \, p(s, b) ds db. \tag{3.2}$$

which can be decomposed into

$$G = U + V$$

where

$$U = \int_0^1 \int_0^1 (t(s, b) - s) \, p(s, b) ds db$$

$$V = \int_0^1 \int_0^1 (b - t(s, b)) \, p(s, b) ds db.$$

are the seller and buyer's expected payoffs, respectively.

Introducing cognitive dispositions during trade, we now assume that each of the parties may be subject to a preference drift in the course of bargaining, manifested by an

---

[6]To ease the exposition, we pursue the analysis with the uniform distribution, though a similar analysis can be carried out also with more general distributions – see Heifetz and Segev (2001) for details.

additive distortion of its valuation. In particular, we assume that the seller's *perceived* valuation of the entitlement consists of $s + \varepsilon$, which represents the sum of her core valuation (s) and a distortion component (ε). Similarly, the buyer's *perceived* valuation of the entitlement consists of $b - \tau$, which represents the difference between his core valuation (b) and a distortion component (τ). Intuitively, ε and τ represent a type of emotional bargaining "toughness" exhibited by each side. Although the seller's "genuine" valuation of the entitlement is s, when bargaining over a sale she becomes convinced that her true valuation is ε dollars higher still. Similarly, the buyer becomes convinced that his valuation is τ dollars lower than his true valuation b.

Such distortions have both empirical and theoretical justifications for coming about. Empirically, there is a vast and growing literature exploring the so-called "endowment effect" in bargaining, which operates much like the toughness distortion envisioned here [see Horowitz and McConnel (2002) or Arlen, Spitzer & Talley (2002) for a literature survey]. Theoretically, the above distortions may play a valuable role in enhancing each side's expected payoff during bargaining. Indeed, if each party to a negotiation perceives herself to possess a more "stingy" valuation than she would possess outside the bargaining context, and this perception is observed by the other bargaining party, then the preference distortion can, ironically, enhance her expected payoff (when viewed from the standpoint of her genuine, a-contextual preferences).

We therefore assume in what follows that the supports of the perceived-valuations – $[\varepsilon, 1 + \varepsilon]$ for the seller and $[-\tau, 1 - \tau]$ for the buyer – are commonly known, but we allow for them to be endogenously determined over time as part of a preference equilibrium. How is this equilibrium determined?

For any *given* toughness dispositions ε, τ, trade takes place with a positive probability only when the true valuation s of the seller is in fact smaller than $1 - \tau - \varepsilon$,[7] and the true valuation b of the buyer is larger than $\tau + \varepsilon$.[8] We assume that the original bargaining mechanism, as characterized by the trade probability $p(s, b)$ and transfer $t(s, b)$ functions, is simply re-scaled to these new intervals of smaller length $1 - \tau - \varepsilon$. The overall probability of trade thus shrinks to

$$\int_0^1 \left(\int_0^1 p(s,b)(1 - \tau - \varepsilon)ds \right) (1 - \tau - \varepsilon)\, db = P(1 - \tau - \varepsilon)^2 \qquad (3.3)$$

and the ex ante gains from trade (as perceived by the bargaining parties) decrease to

$$\int_0^1 \left(\int_0^1 ((1 - \tau - \varepsilon)(b - s))\, p(s,b)\, (1 - \tau - \varepsilon)\, ds \right) (1 - \tau - \varepsilon)db = G(1 - \tau - \varepsilon)^3. \quad (3.4)$$

Of this amount, $U(1 - \tau - \varepsilon)^3$ is enjoyed by the seller and the remaining $V(1 - \tau - \varepsilon)^3$ by the buyer.

---

[7]That is, when her perceived valuation S is smaller than $1 - \tau$, which is the maximal perceived valuation of the buyer.

[8]That is, when his perceived valuation B is higher than ε – the minimal perceived valuation of the seller.

Note, however, that the private payoffs characterized above are expressed in terms of each bargaining party's preferences she perceives herself to have from within the bargaining context. Of equal importance is how these distortions affect the parties "genuine" payoffs away from the bargaining context. Under this metric, the "true" profit of the seller reaps increases by ε for every transaction she successfully consummates (reflecting the abandonment of her transitory cognitive attachment to the entitlement). As such, the seller's a-contextual payoff in the above game in expected value terms is given by:

$$f_{\text{seller}}(\varepsilon, \tau) = U(1 - \tau - \varepsilon)^3 + P(1 - \tau - \varepsilon)^2 \varepsilon \tag{3.5}$$

Similarly, the "genuine" ex ante expected profit of the buyer increases by τ for every successfully-consummated transaction, becoming:

$$f_{\text{buyer}}(\tau, \varepsilon) = V(1 - \tau - \varepsilon)^3 + P(1 - \tau - \varepsilon)^2 \tau. \tag{3.6}$$

Consequently, the a-contextual joint surplus of the parties is given by:

$$g(\varepsilon, \tau) = G(1 - \tau - \varepsilon)^3 + P(1 - \tau - \varepsilon)^2 (\varepsilon + \tau) \tag{3.7}$$

Definition 1 *The bargainers preferences are at equilibrium, if each bargainer's preferences confer the highest expected actual payoff given the preferences of the other bargainer, i.e. if the seller's "endowment effect" ε^* maximizes her expected payoff given the "toughness disposition" τ^* of the buyer, and vice versa.*

In other words, (ε^*, τ^*) are equilibrium dispositions if they constitute a Nash equilibrium of the meta-game with payoffs $f_{\text{seller}}, f_{\text{buyer}}$, which is straightforward to compute:

Proposition 1 *When $\frac{U}{P}, \frac{V}{P} < \frac{1}{3}$, the dispositions with the equilibrium preferences are*

$$\varepsilon^* = \frac{P - 3U}{P + 3(P - G)}$$
$$\tau^* = \frac{P - 3V}{P + 3(P - G)}$$

In fact, a sharper result obtains: In a population of individuals who are repeatedly matched at random to bargain, the preferences will indeed converge to having these levels ε^*, τ^* of the dispositions under any dynamic process that rewards material success with proliferation, for a genuinely wide range of initial distributions of preferences in the population. The proof of this theorem and the exact phrasing and proof of the sharper result requires few more technical definitions, and can be found in Heifetz and Segev (2001).

To grasp the meaning of the conditions $\frac{U}{P}, \frac{V}{P} < \frac{1}{3}$ (and hence $\frac{G}{P} < \frac{2}{3}$), notice that (by (3.1) and (3.2)) it is always the case that $P \leq G$, and the two quantities become

11

closer to one another when the probability of trade $p(s,b)$ decreases when $b-s$ is small, and increases when $b-s$ is large.[9] Therefore, the condition $\frac{G}{P} < \frac{2}{3}$ means that the trade scheme allows for "a fair chance to strike even fairly profitable deals."[10]

Even though a tough spirit or character in the course of bargaining is unilaterally beneficial, and hence both bargainers adapt to such a tough mood during the bargaining process, these tendencies constitutes a Prisoners' Dilemma type of inefficiency:

Proposition 2 *The parties might be better off without the distortion than with it, i.e. there are values for ε and τ such that $G \geq g(\varepsilon, \tau)$. In particular they are better off without the equilibrium distortion than with it, i.e. $G \geq g(\varepsilon^*, \tau^*)$.*

Proof: In the Appendix.

Thus, in terms of their a-contextual preferences (and a fortiori in terms of the traders' preferences in the heat of bargaining), cognitive biases might make both parties worse off than they would be if such biases were nonexistent. It is in precisely such instances that there may be a case for some form of measured paternalism, either by the state or by some other benevolent third party. Mitigating the cognitive shifts during trade is thus a new task for a social planner, on top of the classical task of mitigating strategic misrepresentation. In section 4 we shall explore how various measures of intervention perform in obtaining this duo of goals simultaneously.

3.1 An Example: Sealed-bid Double Auctions

In order to illustrate how the above characterization of equilibrium preferences is operationalized, consider the canonical bargaining problem presented in Chatterjee & Samuelson (1983). Within their model there is a unique Bayes-Nash linear strategy equilibrium profile in which strategies are strictly increasing in type. In the case of equal bargaining power and uniform distributions on $\left[\underline{S}, \overline{S}\right] \times \left[\underline{B}, \overline{B}\right]$, a seller with valuation s offers:

$$\sigma(s) = \frac{2}{3}s + \frac{1}{4}\overline{B} + \frac{1}{12}\underline{S}, \tag{3.8}$$

and a buyer with valuation b bids:

$$\beta(b) = \frac{2}{3}b + \frac{1}{12}\overline{B} + \frac{1}{4}\underline{S}. \tag{3.9}$$

---

[9] $P = G$ only in the limiting case when trade takes place with a positive probability exclusively when $b - s = 1$, i.e. $b = 1$ and $s = 0$.

[10] If we restrict attention to Incentice-Compatibe (IC) and Individually-Rational (IR) trade mechanisms (i.e. Budget-Balanced (BB)), substituting the uniform distributions into inequality (2) of Myerson and Satterthwaite (1983) yields that such mechanisms must satisfy $\frac{G}{P} \geq \frac{1}{2}$. Thus, our condition is compatible with (IC),(IR) and (BB). Moreover, virtually all the particular equilibria of bargaining games we found in the literature satisfy the restriction $\frac{G}{P} < \frac{2}{3}$ (see Heifetz and Segev 2001 for details).

Consequently, trade occurs only when

$$b \geq s + \frac{1}{4}\left(\overline{B} - \underline{S}\right).$$

(3.10)

When normalizing the intervals over the unit square, this condition has the familiar shape

$$b \geq s + \frac{1}{4},$$

and thus trade need not occur even when it is efficient. The total gains from trade in this case are given by:

$$G = \int_0^{\frac{3}{4}} \int_{s+\frac{1}{4}}^1 (b - s)dbds = \frac{9}{64},$$

$U = V = \frac{1}{2}G = \frac{9}{128}$; and finally, the probability of trade is given by:

$$P = \int_0^{\frac{3}{4}} \int_{s+\frac{1}{4}}^1 dbds = \frac{9}{32}$$

The equilibrium biases are therefore[11]

$$\varepsilon^* = \frac{P - 3U}{P + 3(P - G)} = \frac{1}{10}$$

(3.11)

$$\tau^* = \frac{P - 3V}{P + 3(P - G)} = \frac{1}{10}$$

(3.12)

It turns out that with these biases the ex ante probability of trade becomes: $P(1 - \tau^* - \varepsilon^*)^2 = \frac{9}{50}$ which is of course smaller than this probability without the biases (P), and the ex ante total gains from trade are also reduced to $G(1 - \tau^* - \varepsilon^*)^3 + P(1 - \tau^* - \varepsilon^*)^2 (\varepsilon^* + \tau^*) = \frac{27}{250}$. Thus we see that in this case not only that the traders are better off without the distortions than with it (judged by their a-contextual preferences) but also the equilibrium strategies in the presence of the biases induce a less efficient mechanism.

---

[11]To make the computations more explicit, observe that with biases ε, τ, trade can take place only when the *perceived* valuations S, B are in the interval $[\underline{S}, \overline{B}] = [\varepsilon, 1 - \tau]$. Hence, by (3.10), the probability of trade becomes

$$\int_{\underline{S}}^{\overline{B} - \frac{1}{4}(\overline{B} - \underline{S})} \int_{S + \frac{1}{4}(\overline{B} - \underline{S})}^{\overline{B}} dBdS = \int_\varepsilon^{(1-\tau) - \frac{1}{4}(1 - \tau - \varepsilon)} \int_{S + \frac{1}{4}(1 - \tau - \varepsilon)}^{1-\tau} dBdS = \frac{9}{32}(1 - \tau - \varepsilon)^2$$

and the *perceived* gains from trade are

$$\int_\varepsilon^{(1-\tau) - \frac{1}{4}(1 - \tau - \varepsilon)} \int_{S + \frac{1}{4}(1 - \tau - \varepsilon)}^{1-\tau} (B - S)\, dBdS = \frac{9}{64}(1 - \tau - \varepsilon)^3$$

4 Efficient Market Design

Because the existence of endogenous bargaining toughness creates a prima facie case for external intervention, we turn now to exploring the question of what form that intervention might take. As noted in the introduction, we consider three possible candidates, differentiated by the time period in which the social planner enters: intervention ex post, at the interim stage, or at the ex ante stage. For each case, moreover, we consider alternative efficiency definitions using, respectively, the players' a-contextual preferences (which we have labeled "wealth") on the one hand, and their contextualized, "hot" preferences (labeled "happiness") on the other. Although our analysis will focus on the example developed in the previous section, it is easily generalizable.

4.1 Ex post intervention: Subsidizing Trade

It has long been recognized in the bargaining literature that strategic barriers to trade can be reduced – and even eliminated – through an appropriately crafted ex post subsidy. Under such a scheme, a third party insurer promises to pay a subsidy to the bargainers should they successfully consummate a transaction. If the subsidy is sufficiently large, it can counteract the incentives that players might otherwise have to extract information rents by threatening to walk out on the negotiations. The effect can be so pronounced as to eliminate completely the generic inefficiency that frequently attends bilateral bargaining (Myerson & Satterthwaite (1983)).

The attraction to trade subsidies, moreover, is more than a theoretical curiosity. Indeed, a number of legal and institutional mechanisms plausibly serve the very purpose of subsidizing successful bargaining outcomes. While a complete list of them is beyond the scope of this article, a few notable examples are as follows:

- *Tax Incentives:* In state and federal tax law, there are typical deductions and credits that are allowed for certain categories of market purchases.[12]

- *Bankruptcy Costs:* When a firm becomes financially distressed, it is generally agreed that the option of filing for bankruptcy adds considerable costs on the filing party and its creditors (Baird, Jackson & Adler (2000)). The significant costs due to bankruptcy have created substantial motivation for "private workouts" among debtors, their shareholders, and creditors. From a conceptual perspective, then, a successful workout allows the parties to forego a considerable cost, the savings of which can now be split among them. As such, the costs of bankruptcy effectively act as a type of subsidy for successful bargaining.

- *Conditionality in International Aid:* As noted above, numerous donor institutions (e.g., the IMF, World Bank) condition their aid on the resolution of internal or

---

[12]See footnote 3 above.

international conflicts.[13] During the last decade of the 20th century, the amount of international assistance directly tied to the resolution of such conflicts amounted to more than $ 25 billion (Forman and Patrick 2000, p. 10).

- *Anti-Insurance:* Cooter and Porat (2002) have suggested greater use of "anti-insurance" to mitigate incentive problems in joint ventures. Under one example of such a scheme, business partners would execute a contingent debt contract with a third party that would bind the firm to pay off the principal when the firm's profits are low, but excuses the obligation when profits are high. (Because this type of insurance contract increases the volatility of the firm's cash flow, it has been dubbed "anti-insurance"). Although the idea behind anti-insurance is to provide efficient investment incentives on the margin, the same concept might be used to finance an insurance scheme that is triggered with contract negotiations, collective bargaining agreements, or other situations in which bargaining is successful.

Endogenous cognitive dispositions can significantly complicate the considerations underlying an optimal subsidy. Indeed, while continuing to dampen the parties' strategic incentives to misrepresent value, trade subsidies simultaneously raise the absolute size of the bargaining surplus available. This latter effect causes the parties to develop even tougher dispositions than they would have in the absence of subsidies, since a larger surplus enhances the returns that one derives from being committed to a tough mood. Consequently, the optimal subsidy policy will generally have to trade off desirable strategic repercussions with less desirable cognitive ones, and will therefore generally diverge from that implied in a wealth-maximizing actor model.

In order to make the appropriate comparisons, we first consider the optimal trade subsidies in the benchmark case, in which cognitive biases are wholly absent by definition. To focus on intuitions, we restrict attention to the special case in which the expected split of the surplus between the parties is symmetric ($U = V = \frac{G}{2}$). In such a situation, the optimal subsidy scheme generally awards an equal payment to each party upon the consummation of a transaction. Consider, then, the effects of a subsidy α paid to each trader when and only when a trade is consummated. To facilitate welfare comparisons, we shall assume that the cost of the subsidy is wholly internalized ex ante, financed by an ex ante head tax whose size is equal to the expected subsidy paid across all possible valuations.[14]

The inclusion of the subsidy causes the set of mutually advantageous trades to expand by α for each party. Thus, if the interval with gains from trade is originally of length z, its length increases by the total subsidy of 2α to become $z + 2\alpha$. In this case, the optimal

---

[13]Dollar (2000). There is some precedent for this change. Indeed, one of the benefits accrued to Egypt by signing the peace treaty with Israel in 1979 is sustained financial support from the US.

[14]The assumption of self-finance ex ante is not critical. However, regardless of whether the subsidy is self financed or financed by a social insurance scheme, the expected cost of the subsidy is a relevant component of social welfare.

subsidy is that which maximizes the expected total gains of the parties less the cost of the subsidy:[15]

$$\alpha^* = \arg\max_{\alpha} \left[G(1+2\alpha)^3 - 2\alpha P(1+2\alpha)^2 \right] = \frac{1}{6}\frac{3G-P}{P-G} \tag{4.1}$$

and the eventual expected surplus is

$$G(1+2\alpha^*)^3 - 2\alpha^* P(1+2\alpha^*)^2 = \frac{4}{27}\frac{P^3}{(P-G)^2} \tag{4.2}$$

In the numerical example above (where $G = \frac{9}{64}$ and $P = \frac{9}{32}$), this implies that a subsidy of $\alpha^* = \frac{1}{6}$ is required in order to maximize the actual expected surplus. After accounting for their ex ante tax burden, the expected gains from trade will be $G = \frac{1}{6}$, which represents an increase from $G = \frac{9}{64}$ in the absence of the subsidy.

4.1.1 Wealth-Maximizing Subsidies

With this benchmark in hand, we turn to analyze the effects of the endogenous cognitive dispositions. Consider first a social planner whose aim is to craft a subsidy to maximize the expected wealth of the parties. Under this approach, the social planner's problem must now account for the fact that the preference shifts ε, τ, will *generally depend on the subsidy level and eventually adjust to it.* Consequently, an optimal trade subsidy must take this endogeneity into account, as reflected by the following proposition (whose proof can be found in the appendix).

Proposition 3 *With endogenous preferences, a wealth-maximizing social planner will choose a subsidy of*

$$\alpha^{**} = \frac{1}{6}\frac{P}{P-G}$$

*in order to maximize the expected gains from trade, which is **larger** than the optimal subsidy in the benchmark case. The equilibrium dispositions under this subsidy are*

$$\varepsilon^{**} = \tau^{**} = (1+2\alpha^{**})\frac{2P-3G}{2P+6(P-G)},$$

*which are **larger** than those which would emerge without the subsidy. However, the eventual expected surplus will be the same as in the benchmark case with no dispositions —*

$$\frac{4}{27}\frac{P^3}{(P-G)^2}$$

---

[15] In the computations below, the first term is the sum of the ex ante profits of the seller and the buyer, and the second term is the tax.

The intuition behind this result is relatively straightforward. Because the subsidy marginally increases the aggregate bargaining surplus, there is more to be gained for each player from being credibly committed to a tough state of mind. Consequently, distortions in the direction of greater toughness are likely to be increasingly adaptive as the size of the subsidy increases, partially 'cancelling out' the salubrious effects of the subsidy. A social planner must therefore ratchet the subsidy upwards even further to eventually reach a state of first-best efficiency. Once this level of efficiency is attained, however, expected social welfare is identical to that which would emerge under the benchmark case.

In the above example, a subsidy of $\alpha^{**} = \frac{1}{3} > \frac{1}{6} = \alpha^*$ is required to maximize both players' actual ex ante wealth when we take into account their endogenous dispositions. The actual expected surplus upon introducing the subsidy will increase from $\frac{27}{250}$ without the subsidy, to $\frac{1}{6}$ (which is the same as in the benchmark case). The equilibrium dispositions induced by the optimal subsidy grow to $\varepsilon^{**} = \tau^{**} = \frac{1}{6}$ rather than the $\frac{1}{10}$ which would obtain without the external incentive.

4.1.2 Happiness-Maximizing Subsidies

Now consider the alternative case in which the social planner chooses a subsidy in order to maximize the expected sum of the parties' perceived level of happiness at the time of bargaining. This objective corresponds to the following expression.

$$U \left(1 - \varepsilon^{***} - \tau^{***}\right)^3 + V \left(1 - \varepsilon^{***} - \tau^{***}\right)^3$$

where $\varepsilon^{***}, \tau^{***}$ correspond to the equilibrium preferences under the subsidy policy as in proposition 1. As before, we restrict attention to symmetric mechanisms in which $U = V$, so that each level of subsidy influence both players equally. Analysis of this problem generates the following proposition (whose proof can be found in the appendix):

Proposition 4 *With endogenous preferences, a happiness-maximizing social planner will choose a subsidy of*

$$\alpha^{***} = \frac{1}{6} \frac{9G - 4P}{4P - 5G}$$

*which is **smaller** than the optimal subsidy in the benchmark case. The equilibrium dispositions under this subsidy are*

$$\varepsilon^{***} = \tau^{***} = \left(1 + 2\alpha^{***}\right) \frac{2P - 3G}{2P + 6\left(P - G\right)}$$

*which are **larger** than those which would emerge without the subsidy.*

17

The intuition behind this proposition is as follows. Just as in the previous case, provision of a trade subsidy tends to exacerbate the equilibrium level of cognitive dispositions. Unlike that case, however, here the social planner's objective mutates along with the parties' perceived level of happiness at the time of bargaining. Since increasing the bargaining subsidy induces players to perceive that they are tougher bargainers (thereby reducing perceived gains from trade), bargaining failure imposes a smaller social cost on the parties. As a result, the optimal subsidy stops short of that in either the benchmark case or the wealth-maximizing case.

In the running numerical example, maximizing happiness requires imposing a subsidy of $\alpha^{***} = \frac{1}{18}$. This subsidy is clearly smaller than the optimal subsidy $\alpha^* = \frac{1}{6}$ in the benchmark case, and $\alpha^{**} = \frac{1}{3}$ in the wealth-maximization case. The equilibrium dispositions with the optimal subsidy α^{***} will adjust, and increase to $\varepsilon^{***} = \tau^{***} = \frac{1}{9}$ instead of $\frac{1}{10}$ without the subsidy. The expected happiness, given the maximizing subsidy, would rise from $G\left(1 - \tau^* - \varepsilon^*\right)^3 = \frac{9}{125}$ without the subsidy to $\frac{16}{27}\frac{P^3}{(4P-5G)^2} = \frac{2}{27}$ with it.

The following tables synthesize and compare the numerical example in the three cases studied above: the first table compares the benchmark case (no dispositions) and the wealth-maximizing case:

	Benchmark Case No dispositions	Maximizing Wealth
Endogenous disposition without subsidy	0	$\frac{1}{10}$
Expected surplus without subsidy	$\frac{9}{64}$	$\frac{27}{250}$
Optimal subsidy	$\frac{1}{6}$	$\frac{1}{3}$
Endogenous disposition with optimal subsidy	0	$\frac{1}{6}$
Expected surplus with optimal subsidy	$\frac{1}{6}$	$\frac{1}{6}$

The second table compares the benchmark case (no dispositions) and the happiness-maximizing case:

18

	Benchmark Case No dispositions	Maximizing Happiness
Endogenous disposition without subsidy	0	$\frac{1}{10}$
Expected happiness without subsidy	$\frac{9}{64}$	$\frac{9}{125}$
Optimal subsidy	$\frac{1}{6}$	$\frac{1}{18}$
Endogenous disposition with optimal subsidy	0	$\frac{1}{9}$
Expected happiness with optimal subsidy	$\frac{1}{6}$	$\frac{2}{27}$

Note once again from the third row in the tables that the optimal subsidy in the benchmark case systematically diverges from that in either of the other two cases involving endogenous bias, falling short of the wealth-maximizing subsidy and exceeding the happiness-maximizing subsidy. Moreover, note from the fourth row in the tables that the equilibrium level of predicted toughness is not uniform across the two alternative objectives, and is significantly higher in the case of maximizing wealth. These respective differences exemplify our more general argument in this paper: That accounting for the endogenous effects of regulation itself can lead to predictions that are distinct from those that would be rendered if one either ignored cognitive biases or treated them as an exogenous primitive.[16]

4.2 Interim Intervention: Efficient Bargaining Mechanisms

Another important arena for regulatory intervention comes at the interim stage, in the design of bargaining procedures themselves. It is widely recognized in the literature that bargaining protocols matter, in that they can produce distinct trading probabilities and expected social payoffs. Consequently, the question of what constitutes an "optimal bargaining mechanism" in a given circumstance continues to receive a significant amount of attention. At the very least, the features of an optimal mechanism identify the limits of what can be accomplished in an un-mediated bargaining (Fudenberg and Tirole (1991, p. 290)). In this section, then, we explore optimal bargaining mechanisms, taking into account endogenous dispositions.

At core, all bargaining mechanism specify both (1) the probability of trade for any pair of valuations, and (2) the distribution of gains from trade that ensue from a transaction.

---

[16]Note also, by comparing the first and third entries of the last line of the table, that if both bargainers were cold-blooded and incapable of developing cognitive distortions in the bargaining phase, they would both be happier *even under the best-suited subsidy intervention.* Thus, even though these cognitive distortions may be unilaterally beneficial (when the preferences of the other bargainer are taken as given), their concurrent presence is socially destructive, just as in a Prisoners' Dilemma. The institutional intervention can amend this state of affairs, but to a limited extent.

In conventional models no such mechanism can be considered optimal if there is an alternative incentive compatible, individually rational mechanism that produces trade in strictly more situations. As we shall see below, however, the introduction of endogenous dispositions can sometimes cause the optimal mechanism to diverge from this general principle. In particular, under certain conditions, an optimal mechanism may be much more "draconian" than theory would otherwise predict, refusing to enforce transactions in strictly *fewer* situations than other implementable mechanisms.

Interestingly, certain well-known doctrines operate in much the same fashion, and can be interpreted (at least indirectly) as requiring some artificial lower bound in trade surplus before a court is willing to enforce a contract. Notably, most of the protections are limited to special cases, and contract law doctrine more generally is thought to implement the principle that courts should act to facilitate transactions in the most circumstances possible.

- *Unconscionability:* This doctrine instructs the court to refuse to enforce contracts in situations where the negotiation and resulting terms of a transaction are excessively one sided.[17] One interpretation of the unconscionability doctrine is that it essentially requires that each party receives some minimum surplus before a contract can be enforced. Under such an interpretation, the doctrine implies a requirement that the total amount of social surplus exceed some specified threshold before a contract is enforceable.

- *Moieties:* In admiralty law, the common-law doctrine of moieties dictated the division of rents from emergency salvage operations at sea. When, for instance, a vessel in distress off-loaded its cargo onto another ship that had come to its aid, the doctrine required that each party was to receive a "moiety" of a fixed fraction (usually either $\frac{1}{3}$ or $\frac{1}{2}$) of the value of the cargo as computed by its trading price once sold on a commodities market.[18] Because the moieties doctrine constrains the feasible set of negotiated outcomes by fixing a price by reference to an external market, it has the likely effect of discouraging transactions in the bilateral monopoly setting where the gains from trade are positive, but insubstantial.

- *Cooling off periods:* A number of state and federal laws in the United States require a specified period of time to pass before certain types of consumer contracts are enforceable. In federal law, for example, statutory cooling off periods are required for door-to-door sales (16 C.F.R. Part 429), telemarketing, (16 C.F.R. Parts 308,

---

[17]See, e.g., Williams v. Walker Thomas Furniture Co. 350 F.2d 445 (D.C. Cir. 1965) (refusing enforcement of a cross-collateralization clause in a consumer debt contract that would operate to preclude satisfaction on the debt of any single purchase until all purchases were paid off).

[18]See, e.g., Post v. Jones, 60 U.S. 150 (1856), in which a distressed vessel actually implemented a competitive bidding process among three aspiring salvagers. The lowest bidder later successfully challenged the terms of the contract under the doctrine of moeities, arguing that its accepted bid was too low to ensure it of receiving its equitable share of the rents under the doctrine.

310, & 435), and sales of business franchises (16 C.F.R. Secs. 436.1 et seq.). As with unconscionability doctrine, one interpretation of a cooling off period is as a doctrine that requires gains from trade to be relatively large. Indeed, where the gains are small at the time of negotiation, it is relatively likely that small post-transaction perturbations of the parties' respective valuations can cause recision of the agreement within the cooling off period. Knowing this, rational parties might not find it in their interests to consummate deals that are likely to prove unenforceable.

Given these examples, we now turn to consider the characteristics of an optimal mechanism under endogenous biases. Such a mechanism is characterized by the pair of functions $(p(s,b),t(s,b))$ where $p(s,b)$ and $t(s,b)$ denote the equilibrium probability of trade and the equilibrium transfer payment, respectively, between a seller with a valuation s, and a buyer with a valuation b.

Once again, as a benchmark we first specify the mechanism which maximizes the players payoff assuming away the emergence of biases. As demonstrated by Myerson & Satterthwaite (1983), the most efficient Incentive-Compatible (IC), Individually-Rational (IR) and Budget-Balanced (BB) mechanism is one in which trade is prohibited unless the seller's and buyer's stated valuations differ by at least $\frac{1}{4}$, i.e. a mechanism in such that

$$p(s,b) = \begin{cases} 1 & b-s \geq \frac{1}{4} \\ 0 & b-s < \frac{1}{4} \end{cases} .$$

The mechanism presented in the example in section 2.2 is thus an optimal one. The expected payoffs U, V of the bargainers might be different from one efficient mechanism to another. On the other hand, in every such optimal mechanism the total surplus and the expected probability of trade remain the same –

$$G = \int_0^1 \int_0^1 (b-s)\,p(s,b)dsdb = \int_{\frac{1}{4}}^1 \int_0^{b-\frac{1}{4}} (b-s)\,dsdb = \frac{9}{64}$$

$$P = \int_0^1 \int_0^1 p(s,b)dsdb = \int_{\frac{1}{4}}^1 \int_0^{b-\frac{1}{4}} dsdb = \frac{9}{32}$$

4.2.1 Wealth-Maximizing Mechanisms

As before, we now consider how the introduction of endogenous dispositions affects the analysis, focusing first on the social objective of maximizing the expected ex ante wealth of the bargainers. Also as before, to facilitate intuitions, we restrict attention to the symmetric case where $U = V = \frac{1}{2}G$. Under this formulation, the expected sum of seller and buyer welfare is:

$$U\left(1 - \tau^* - \varepsilon^*\right)^3 + P\left(1 - \tau^* - \varepsilon^*\right)^2 \varepsilon^*$$

21

where ε^* and τ^* are the equilibrium biases from proposition 1. Analysis of the bargaining design problem gives rise to the following proposition (whose proof appears in the appendix):

Proposition 5 *A Myerson-Satterthwaite (1983) mechanism with*

$$p(s,b) = \begin{cases} 1 & b - s \geq \frac{1}{4} \\ 0 & b - s < \frac{1}{4} \end{cases}$$

maximizes the expected wealth of the traders among all IC, IR and BB mechanisms, even when the endogenous biases ε^*, τ^* from Proposition 1 are taken into account. When the expected gains from trade are shared equally, the equilibrium dispositions are $\varepsilon^* = \tau^* = \frac{1}{10}$.

Interestingly, when the social planner is motivated by maximizing wealth, the efficient mechanism permits trade only when the seller's and buyer's reported valuations differ by at least $\frac{1}{4}$, exactly as in Myerson-Satterthwaite (1983).

Consequently every efficient mechanism as in Myerson-Satterthwaite is also efficient in order to maximize wealth paternalistically among distorted bargainers.

4.2.2 Happiness-Maximizing Mechanisms

Suppose instead that the social planner were motivated by a desire to maximize happiness rather than wealth. Under this alternative objective, the planner's maximand becomes:

$$U \left(1 - \varepsilon^{***} - \tau^{***}\right)^3$$

Where ε^{***} and τ^{***} are the equilibrium biases from proposition 1. Analysis of this problem yields the following proposition (whose proof appears in the appendix):

Proposition 6 *The mechanism that maximizes the happiness of the players is characterized by a threshold $h^{***} = \frac{1}{2}$ such that*

$$p(s,b) = \begin{cases} 1 & b - s \geq 1 - h^{***} = \frac{1}{2} \\ 0 & b - s < 1 - h^{***} = \frac{1}{2} \end{cases}$$

Such a mechanism induces trade in strictly fewer instances than the optimal mechanism in the baseline case. The equilibrium dispositions under this mechanism are $\varepsilon^{***} = \tau^{***} = 0$.

Note that the optimal happiness-maximizing mechanism is a significantly more "draconian" bargaining mechanism than that of either the baseline case or of the wealth-maximizing, allowing trade if and only if the gains from trade $(b - s)$ exceed $\frac{1}{2}$. The intuition behind this result stems from a fundamental tradeoff that a restrictive mechanism creates. On the one hand, more restrictive trading mechanisms impose a direct

welfare loss, since they reduce the likelihood of any trades. On the other hand, this reduction in the likelihood of trade reduces the adaptiveness of a tough bargaining strategy, since the size of the expected bargaining surplus is smaller. The result reported in Proposition 6 reflects the fact that the latter effect swamps the former one for all positive dispositions, so that an optimal trading rule coincides with the least restrictive mechanism that completely vitiates all biases. Consequently, under this mechanism, no biases ever evolve. It is easily verified that this mechanism yields a probability of trade of $P = \frac{1}{8}$, and an average payoff per player of $U = \frac{1}{24}$.

Note also that unlike the previous examples, the happiness-maximizing mechanism here is Pareto inferior to other candidates. For example, the Myerson-Satterthwaite (1983) optimal mechanism is clearly implementable in the case of zero biases, and both parties would prefer its implementation to the one given in the proposition. Allowing them to do so, however, would cause the parties to evolve increasingly tough dispositions, which in the long run would yield less trade and less ultimate happiness (as evaluated at the time of trade). Consequently, implementing the happiness maximizing mechanism would require courts to actively prohibit trade except in situations where the surplus is sufficiently high. Many of the immutable legal doctrines discussed at the beginning of this section attempt to do just that.

Finally, note that just as with subsidies, the maximand favored by the social planner has a clear effect on the ultimate allocational rule. Maximizing wealth and maximizing happiness lead to very different solutions.

4.3   Ex ante intervention: Property Rights

For our final application, we consider regulatory interventions that occur before bargaining ever begins. Because transactions are little more than the transfer of property rights, it can be substantially affected by calculated manipulations to the content of those initial property rights. Indeed, the use of divided property rights has already been cited as a way to address problems of *strategic* barriers to trade (e.g., Cramton, Gibbons & Klemperer (1987); Ayres & Talley (1995)).

Divided entitlements might, on first blush, appear exceptional within a capitalist economy. But on closer inspection, one can find dozens of areas where either courts or the parties themselves provide for divided property rights. Although a complete description of such partial entitlements is too lengthy to articulate here, the following represents a reasonable sampling:

- *Outright Co-ownership:* In trade secret law, the "shop rights" doctrine provides for a type of divided ownership of inventions developed in the workplace. Explicitly, when a rank-and-file employee uses company time and/or resources in developing a new invention, the employee is awarded general ownership rights to the invention, while the employer receives a non-exclusive, zero-price license to use the invention.

Employers and employees are generally free to contract around the shop rights doctrine, either prior to or after invention occurs (See Lester & Talley, 2000).

- *Temporal/Subject Matter Divisions:* In patent law, patentees are generally awarded with a strong property right, but one that runs for only a prescribed, 20-year statutory period after the effective filing date (or 17 years from the date of the grant) (See 35 U.S.C. §§ 119, 120, 154(a)(2) (2002)). Viewed ex ante, this prospective temporal division can be thought to convey payoffs whose present value is divided between the patentee and its competitors who wish to use the patented technology. Once again, patent law allows for contracting around this statutory entitlement through licensing agreements.

- *Legal Uncertainty:* In business law, corporate fiduciaries are prohibited from appropriating "corporate opportunities" — i.e., prospective business ventures that rightfully belong to the firm — for their own personal use. The standards that identify what exactly constitutes a bona fide corporate opportunity, however, are inherently casuistic, leaving an obscure doctrine that has been alternatively characterized by legal commentators as "vague," "in transition," "far from crystal clear," and "indecipherable" (see Talley (1998)). Although such randomness is generally perceived as undesirable, it has the effect of endowing both the fiduciary and the corporation with a probabilistic claim on the business opportunity, which has a number of characteristics resembling joint ownership. Moreover, corporate law generally allows for firms and their fiduciaries to allocate opportunities through bargaining.

- *Liability Rules:* Even in the absence of a physical, temporal or probabilistic division of property, legal rules can divide claims by altering the form of protection accorded one's entitlement. Much of modern nuisance law in the United States, for example, tends to award a successful plaintiff with money damages rather than injunctive relief for a defendant's incompatible activities. As with the above examples, the plaintiff and defendant are free to negotiate in the shadow of this liability rule (see Ayres & Talley (1995)).

In order to consider the effect of divided entitlements on bargaining with dispositions, suppose the parties bargained over ownership of an asset, but now assume that the initial property rights of the asset are given by $(q, 1 - q)$, where q represents the fractional ownership claimed by B, and $(1 - q)$ represent that claimed by S. Without loss of generality, suppose that $q \leq \frac{1}{2}$. Following Ayres & Talley (1995) we explore what is the optimal allocation (i.e. what is the optimal q) that maximizes the expected surplus in case the partners wish to dissolve the partnership, but we now factor in the possibility of endogenous cognitive dispositions.

As a benchmark, we once again consider first the case in which no biases exist, and we explore a specific bargaining procedure: a sealed-bid double auction as the procedure to

dissolve the partnership: Each of the partners submits a bid for the asset, and the partner with the higher bid buys her partner's share in the asset at the price (for the entire asset) which is the average of the two bids. Such a procedure with a linear-strategy equilibrium is known to be optimal in a Myerson-Satterthwaite framework with symmetric, uniform distributions of valuation.

As in section 3.1, we therefore explore an equilibrium with linear bidding strategies. These turn to be

$$r_B(b) = \begin{cases} \frac{2}{3}b + \frac{1}{12} + \frac{1}{6}q & \text{when } \frac{1}{4} - \frac{1}{2}q \leq b \leq 1 \\ \frac{1}{4} - \frac{1}{6}q & \text{when } 0 \leq b < \frac{1}{4} - \frac{1}{2}q \end{cases} \qquad (4.3)$$

$$r_S(s) = \begin{cases} \frac{2}{3}s + \frac{1}{4} - \frac{1}{6}q & \text{when } 0 \leq s \leq \frac{3}{4} + \frac{1}{2}q \\ \frac{3}{4} + \frac{1}{6}q & \text{when } \frac{3}{4} + \frac{1}{2}q < s \leq 1 \end{cases} \qquad (4.4)$$

(of which the Chatterjee-Samuelson (1983) equilibrium in section 3.1 is the limiting case for $q = 0$[19]). Analysis of these expressions leads to the following proposition for the benchmark case of no dispositions (whose proof appears in the appendix):

Proposition 7 *In the absence of dispositions, the expected surplus in this double auction equilibrium is maximized with the initial shares $(q, 1-q) = \left(\frac{1}{2}, \frac{1}{2}\right)$. The asset will always end up in the hands of the partner who values it most.*

The fact that the optimal property rights scheme in the baseline case allocates equal ownership shares to each player should not be surprising. Indeed, in this case, the only impediment to a negotiated outcome is the parties' incentives to extract information rents by misstating their true valuations. Buyers tend to shade their private valuations downwards, while sellers tend to shade theirs upwards. A division of property rights tends to weaken these incentives, by making each player both a potential buyer and a potential seller. The introduction of these dual roles causes the players to become more ambivalent about whether to shade their valuations (and in which direction to do so). When each side has a one-half initial ownership share of the asset, the incentives to over-state and under-state exactly cancel one another out, thereby leading to first-best efficiency.

4.3.1 Wealth-Maximizing Property Rights

We now consider how a social planner might maximize surplus in the presence of endogenous cognitive dispositions. To understand how biases alter the analysis, first fix q

---

[19]The only difference is that in the Chaterjee-Samuelson (1983) case ($q = 0$), no trade takes place when $0 \leq b < \frac{1}{4} - \frac{1}{2}q$ or $\frac{3}{4} + \frac{1}{2}q < s \leq 1$ with either the equilibrium strategies (3.3), (3.4) or alternatively (4.1), (4.2). However, when $q > 0$ and $0 \leq b < \frac{1}{4} - \frac{1}{2}q$, the partner with share q is certain to sell its part in the partnership at equilibrium, and will therefore bid the lowest equilibrium bid of its partner, and not below it. Similarly, when $\frac{3}{4} + \frac{1}{2}q < s \leq 1$, the partner with share $1 - q$ is certain to buy its partner's part, and will therefore bid the highest equilibrium bid of its partner, and not above it.

and consider only how dispositions are likely to evolve. As above, suppose that player S misperceives her valuation to be ε higher than it actually is, while B similarly misperceives his valuation to be τ lower than it actually is. Consequently, when the partners begin to negotiate, they observe each other's character (i.e. the supports $[\varepsilon, 1 + \varepsilon]$ and $[-\tau, 1 - \tau]$ of the distributions become common knowledge), but not the actual perceived valuations. Then they play the equilibrium where the bids r_B, r_S are linear in their perceived valuations. When translated back to their actual valuations, these equilibrium bids turn to be

$$
r_B(b) = \begin{cases} \frac{2}{3}b + \frac{1}{12} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau & \text{when } x \leq b \leq 1 \\ \frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau & \text{when } 0 \leq b < x \end{cases}
\tag{4.5}
$$

$$
r_S(s) = \begin{cases} \frac{2}{3}s + \frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau & \text{when } 0 \leq s \leq 1 - x \\ \frac{3}{4} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau & \text{when } 1 - x < s \leq 1 \end{cases}
\tag{4.6}
$$

where

$$
x = \frac{1}{4} - \frac{1}{2}q + \frac{3}{4}(\varepsilon + \tau).
\tag{4.7}
$$

Analysis of these bid functions yields the following proposition (whose proof can be found in the appendix):

Proposition 8 *In the double auction with a given value of q, the equilibrium dispositions of the parties are*

$$
\varepsilon^* = \tau^* = \frac{3}{10} + \frac{1}{3}q - \frac{1}{15}\sqrt{(9 + 80q)}.
$$

Expected wealth is maximized when the initial entitlements are fixed at $(q, 1 - q) = \left(\frac{1}{2}, \frac{1}{2}\right)$, so that

$$
\varepsilon^* = \tau^* = 0
$$

Proposition 7 illustrates that a wealth-maximizing property rights division under endogenous cognitive biases is identical to that in the baseline case: each party receives a one-half ownership share in the asset. This result suggests that there may be at least some forms of regulatory intervention that can address both strategic barriers to trade and cognitive barriers to trade simultaneously. Indeed, not only does divided ownership dampen information rents (as is well known in the literature), but Proposition 7 demonstrates that it can also dampen the returns to developing a toughness disposition. When the parties could ultimately be either buyers or sellers of the asset, there is little to be gained from being committed to a set of preferences that biases one's valuation either upwards or downwards. Consequently, the allocation of equal shares $(q, 1 - q) = \left(\frac{1}{2}, \frac{1}{2}\right)$ both maximizes the players wealth and also induces an equilibrium with no cognitive dispositions.

4.3.2 Happiness-Maximizing Property Rights

The intuition developed in Proposition 7 turns out to be quite general. In fact, not only does an initial property right division maximize expected wealth, but it also turns out to be the optimal property rights allocation when the social planner's objective is to maximize happiness, as reflected in the following proposition (whose proof is in the appendix):

Proposition 9 *In the double auction with a given value of q, expected happiness is maximized when the initial entitlements are fixed at $(q, 1 - q) = (\frac{1}{2}, \frac{1}{2})$, so that*

$$\varepsilon^* = \tau^* = 0$$

Recall that in the ex-post and interim cases (studied above), the optimal regulatory intervention hinged crucially on the objective function of the social planner, and in particular whether she was motivated by maximizing "wealth" or "hapiness." Here, in contrast, an ex ante intervention towards evenly divided property rights turns out to be optimal under both criteria as well as under the baseline case. Moreover, such a division has the effect of completely debiasing the players, so that they no longer develop cognitive distortions in the course of bargaining. This prediction squares nicely with some recent experimental work (e.g., Rachlinsky & Journden (1998)), in which endowment effects appear to dissipate when parties' interests are protected by weaker entitlements (such as liability rules).

5 Conclusion

This paper has presented a framework for designing optimal institutions in the presence of endogenous cognitive dispositions. This is a critically important problem if one wishes to draw meaningful policy implications from behavioral economics. At the same time, however, it is a problem that involves at least two unique complicating factors which are largely absent in conventional institutional design problems. First, the existence and size of cognitive biases may themselves be sensitive to the very institutional policies designed to address them. In such situations, policymakers must be keenly aware of the feedback effects that any candidate mechanism is likely to foster, and anticipate how cognition and regulatory context are likely to interact. Second, the very definition of "optimality" may be even more contestable when preference are endogenous. Policies designed to maximize wealth (i.e., welfare defined in terms of a-contextual preferences) need not coincide with those designed to maximize happiness (i.e., welfare defined in terms of the preferences induced within the institutional design). Consequently, policymakers may be forced to choose between these alternative objectives, since they generally will not produce the same policy prescriptions.

To illustrate our claims, we considered three families of regulatory intervention that have real-world institutional counterparts: ex post, interim and ex-ante interventions in bilateral trade. Within these contexts, we have demonstrated how a failure to appreciate the complicating factors noted above can lead to unintended and undesirable institutional structures. In the context of ex post intervention, the optimal trade subsidy that incorporates evolutionary biases always diverges from the "baseline" case in which biases are ignored, regardless of the social objective adopted. In particular, wealth maximization requires a larger subsidy relative to the baseline, while happiness maximization requires a smaller one. Moreover we have shown that an ex post intervention (of any size) induces the players to have even larger perception biases.

For interim interventions, the optimal trading mechanism in the baseline case turns out to be identical to that of a wealth-maximizing mechanism with biases. On the other hand, the trading mechanism that maximizes the players' happiness turns out to be relatively "draconian" in nature, prohibiting trade in strictly more circumstances than other implementable bargaining mechanisms would. Implementing such a Pareto-inferior mechanism would likely necessitate the implementation of immutable rules (such as that found in the doctrines of unconscionability or moeities).

Most optimistically, we find that ex ante regulation through property rights allocations may be the most flexible and promising of all the interventions studied (at least within our framework). Here, the optimal allocation entails divided ownership, awarding half of the entitlement to each player. This allocation remains optimal regardless of whether the objective is to maximize wealth or happiness, and of whether we take the biases into account. Moreover, such a regulatory scheme completely de-biases the players, eventually eliminating their evolutionary dispositions.

Although it is our motivating story, bilateral trade is likely not to be the sole arena in which the interplay between policy, context and preferences is important. There are already several other attempts to address similar issues in other contexts.[20] Indeed, the approach suggested here may be relevant in virtually every instance of market or institutional design, and hence suggests a promising direction for further research.

6 Appendix

Proof of Proposition 2:

It is enough to prove the second part of the proposition. If the dispositions are indeed the equilibrium dispositions we have

$$g\left(\varepsilon^{*}, \tau^{*}\right) = \frac{4P^{3}(2P - G)}{(4P - 3G)^{3}}$$

From proposition 1 we have the conditions $\frac{P}{U} > 3$ and $\frac{P}{V} > 3$ and thus $\frac{P}{G} > \frac{3}{2}$. We also assume $\frac{P}{G} \leq 2$ since we restrict our attention to mechanisms which are both (IC) and (IR) (see footnote 10). Finally, $g\left(\varepsilon^*, \tau^*\right)$ is indeed smaller than G for $\frac{3}{2} < \frac{P}{G} \leq 2$. $\blacksquare$

Proof of Proposition 3:

In the case of a subsidy α the equilibrium dispositions for the parties are given by

$$\varepsilon_\alpha = (1 + 2\alpha) \frac{P - 3U}{P + 3\left(P - G\right)} \tag{A.1}$$

for the seller, and

$$\tau_\alpha = (1 + 2\alpha) \frac{P - 3V}{P + 3\left(P - G\right)} \tag{A.2}$$

for the buyer: From Proposition 1, the endogenous toughness is of size $\varepsilon^* = \frac{P - 3U}{P + 3(P - G)}$ and $\tau^* = \frac{P - 3V}{P + 3(P - G)}$ of the length of the interval, which is now $1 + 2\alpha$ instead of 1.

Thus the actual bargaining interval is of length

$$L_{sub} = (1 + 2\alpha - \varepsilon_\alpha - \tau_\alpha) = \left(1 - \frac{2P - 3G}{P + 3\left(P - G\right)}\right)(1 + 2\alpha) \tag{A.3}$$

The actual expected wealth is

$$f_{\text{seller}} = \left[U\left(L_{sub}\right)^3 + P\left(L_{sub}\right)^2 \varepsilon_\alpha\right] - P\left(L_{sub}\right)^2 \alpha \tag{A.5}$$

for the seller and

$$f_{\text{buyer}} = \left[V\left(L_{sub}\right)^3 + P\left(L_{sub}\right)^2 \tau_\alpha\right] - P\left(L_{sub}\right)^2 \alpha \tag{A.6}$$

for the buyer: The first term in f_{seller} and f_{buyer} is computed as in (3.5) and (3.6), respectively; and the second term is the per-capita tax that each of the traders has to pay in advance in order to finance the expected subsidy. The wealth-maximizing subsidy is therefore

$$\alpha^{**} = \arg\max_\alpha \left[G\left(L_{sub}\right)^3 + P\left(L_{sub}\right)^2 \left(\varepsilon_\alpha + \tau_\alpha\right) - 2P\left(L_{sub}\right)^2 \alpha\right] = \frac{1}{6} \frac{P}{P - G} \tag{A.7}$$

Since we assume throughout that $\frac{P}{G} > \frac{3}{2}$ then indeed

$$\alpha^{**} \geq \alpha^* = \frac{1}{6} \frac{3G - P}{P - G}$$

The resulting dispositions $\varepsilon^{**} = \tau^{**}$ will be larger by a factor of $(1 + 2\alpha^{**})$ relative to the dispositions $\alpha^* = \tau^*$ without the subsidy. Finally, the eventual expected surplus is

$$G\left(L_{sub}\right)^3 + P\left(L_{sub}\right)^2 \left(\varepsilon^{**} + \tau^{**}\right) - 2P\left(L_{sub}\right)^2 \alpha^{**} = \frac{4P^3}{27\left(G - P\right)^2},$$

the same as (4.2) in the benchmark case. ∎

Proof of Proposition 4:

The computation of the dispositions $\varepsilon_\alpha, \tau_\alpha$ and the relevant interval size L_{sub} is the same as in (A.1), (A.2) and (A.3) above. However, when maximizing the expected happiness of the traders during trade rather than their expected wealth, the term $P(L_{sub})^2(\varepsilon_\alpha + \tau_\alpha)$ in (A.7) should be omitted from the maximand, since this is the extra expected gains from trade that the biased bargainers do not see in the heat of bargaining. The optimal subsidy is therefore

$$\alpha^{***} = \arg\max_\alpha \left[G(L_{sub})^3 - 2P(L_{sub})^2 \alpha \right] = \frac{1}{6}\frac{9G - 4P}{4P - 5G} \tag{A.8}$$

which is indeed smaller than the optimal subsidy α^* in (4.2) in the benchmark case for the relevant range $\frac{P}{G} > \frac{3}{2}$. ∎

Proof of Proposition 5:

Denote $x = \frac{2P}{G}$. From proposition 1 we have

$$\varepsilon^* = \tau^* = \frac{x - 3}{x + 3(x - 2)}$$

From (3.7) the expected surplus is

$$g(\varepsilon^*, \tau^*) = G\left(1 - 2\frac{x-3}{x+3(x-2)}\right)^3 + 2P\left(1 - 2\frac{x-3}{x+3(x-2)}\right)^2\left(\frac{x-3}{x+3(x-2)}\right) = G\frac{x^3(x-1)}{4(2x-3)^3} \tag{A.9}$$

In the relevant range in which $\frac{P}{G} \geq \frac{3}{2}$, and the trade mechanism is also (IC), (IR) and (BB) (see footnote 10) and hence $\frac{P}{G} \leq 2$, we have $x = \frac{2P}{G} \in [3, 4]$. In this range $\frac{x^3(x-1)}{4(2x-3)^3}$ is a decreasing function of x.

Let

$$G^* = \int_0^1 \int_0^1 \left(\int_0^1 (b-s)\, p^*(s,b)\, ds\, db \right)$$

be the average surplus in the wealth-maximizing mechanism when normalized back to the $[0,1] \times [0,1]$ domain, and

$$P^* = \int_0^1 \int_0^1 p^*(s,b)\, ds\, db$$

be the average probability of trade in that mechanism. From (A.9) we know that among all mechanisms with an average surplus G^*, in the wealth-maximizing one the expression $\frac{x^3(x-1)}{4(2x-3)^3}$ is maximal. Since this expression is decreasing, in the wealth-maximizing

30

mechanism $x = \frac{2P^*}{G^*}$ is minimal in the range $[3, 4]$. We thus have to solve the problem

$$\min_{p:[0,1]\times[0,1]\to[0,1]} \left[\int_0^1 \int_0^1 p(b, s) ds db \right]$$

$$\text{s.t.} \quad \int_0^1 \int_0^1 (b - s) \, p(b, s) ds db = G^*$$

The Lagrangian is

$$L = \int_0^1 \int_0^1 [(1 - \mu(b - s)) \, p(b, s)] ds db$$

and thus the first order condition gives

$$p^*(b, s) = \begin{cases} 1 & b - s \geq \frac{1}{\mu} \\ 0 & b - s < \frac{1}{\mu} \end{cases}$$

This means that there is a threshold $h^* \in [0, 1]$ for which

$$p^*(s, b) = \begin{cases} 1 & b - s \geq 1 - h^* \\ 0 & b - s < 1 - h^* \end{cases}$$

This implies that

$$G^* = \int_0^{h^*} \left(\int_{s+1-h^*}^1 (b - s) \, db \right) ds = \frac{1}{6} h^{*2} \left(3 - 2h^* \right),$$

$$P^* = \int_0^{h^*} \left(\int_{s+1-h^*}^1 db \right) ds = \frac{1}{2} h^{*2}$$

and

$$x^* = \frac{2P^*}{G^*} = \frac{6}{(3 - 2h^*)}.$$

Since $x^* \in [3, 4]$ we conclude that $h^* \in \left[\frac{1}{2}, \frac{3}{4} \right]$. The optimal threshold h^* thus satisfies

$$h^* = \arg \max_{h \in \left[\frac{1}{2}, \frac{3}{4} \right]} \left[G \frac{x^3 (x - 1)}{4 (2x - 3)^3} \right] = \arg \max_{h \in \left[\frac{1}{2}, \frac{3}{4} \right]} \left[\frac{1}{3} (3 + 2h) \frac{h^2}{(1 + 2h)^3} \right] = \frac{3}{4}$$

∎

Proof of Proposition 6:

The expected happiness of the traders during trade is the first term in (A.9)

$$G \left(1 - 2 \frac{x - 3}{x + 3(x - 2)} \right)^3 = G \left(\frac{x}{2x - 3} \right)^3$$

31

and $\left(\frac{x}{2x-3}\right)^3$ is a decreasing function in the relevant range $x \in [3,4]$. Thus the same argument as in the proof of proposition 5 suggests that the happiness-maximizing mechanism would also be "draconian", allowing for trade only if the gains from trade exceed a threshold $1 - h^{**}$, where

$$h^{**} = \arg \max_{h \in \left[\frac{1}{2}, \frac{3}{4}\right]} \left[G\left(\frac{x}{2x-3}\right)^3 \right] = \arg \max_{h \in \left[\frac{1}{2}, \frac{3}{4}\right]} \left[\frac{2}{3}(3 - 2h) \frac{h^2}{(1+2h)^3} \right] = \frac{1}{2}$$

∎

Proof of Proposition 7:

Whenever the bidding is not extreme (i.e. $\frac{1}{4} - \frac{1}{2}q \le b \le 1$ and $0 \le s \le \frac{3}{4} + \frac{1}{2}q$), the trading price for the entire asset is

$$p(s,b) = \frac{r_S(s,b) + r_B(s,b)}{2} = \frac{1}{3}(s+b) + \frac{1}{6}$$

By (4.3) and (4.4). For $q \in [0, \frac{1}{2}]$, the overall expected payoffs of each of the players is the sum of the expected payoffs in 5 mutually exclusive and exhaustive regions:

1) When B gets the entitlement i.e. $r_B > r_S$ (this can only happen when the bids are not extreme i.e. $\frac{1}{4} - \frac{1}{2}q \le b \le 1$, $0 \le s \le \frac{3}{4} + \frac{1}{2}q$ and $b - s > \frac{1}{4} - \frac{1}{2}q$)

2) When S gets the entitlement (i.e. $r_S \ge r_B$) and

2a) Bids are not extreme i.e. $\frac{1}{4} - \frac{1}{2}q \le b \le 1$, $0 \le s \le \frac{3}{4} + \frac{1}{2}q$ and $r_S \ge r_B \Leftrightarrow b - s \le \frac{1}{4} - \frac{1}{2}q$

2b) Only the bid of B is extreme i.e. $0 \le b < \frac{1}{4} - \frac{1}{2}q$ and $0 \le s \le \frac{3}{4} + \frac{1}{2}q$

2c) Both bids are extreme i.e. $0 \le b < \frac{1}{4} - \frac{1}{2}q$ and $\frac{3}{4} + \frac{1}{2}q < s \le 1$

2d) Only the bid of S is extreme i.e. $\frac{1}{4} - \frac{1}{2}q \le b \le 1$ and $\frac{3}{4} + \frac{1}{2}q < s \le 1$

Denote by π_B, π_S the expected payoffs of B and S in this equilibrium.

$$\pi_S = \int_0^{\frac{3}{4}+\frac{1}{2}q} \int_{s+\frac{1}{4}-\frac{1}{2}q}^1 ((1-q)(p(s,b) - s))\, db\, ds + \int_0^{\frac{3}{4}+\frac{1}{2}q} \int_{\frac{1}{4}-\frac{1}{2}q}^{s+\frac{1}{4}-\frac{1}{2}q} q\,(s - p(s,b))\, db\, ds$$

$$+ \int_0^{\frac{3}{4}+\frac{1}{2}q} \int_0^{\frac{1}{4}-\frac{1}{2}q} q\left(s - \frac{1}{2}\left(\left(\frac{1}{4} - \frac{1}{6}q\right) + \left(\frac{2}{3}s + \frac{1}{4} - \frac{1}{6}q\right)\right)\right) db\, ds$$

$$+ \int_{\frac{3}{4}+\frac{1}{2}q}^1 \int_0^{\frac{1}{4}-\frac{1}{2}q} q\left(s - \frac{1}{2}\left(\left(\frac{1}{4} - \frac{1}{6}q\right) + \left(\frac{3}{4} + \frac{1}{6}q\right)\right)\right) db\, ds$$

$$+ \int_{\frac{3}{4}+\frac{1}{2}q}^1 \int_{\frac{1}{4}-\frac{1}{2}q}^1 q\left(s - \frac{1}{2}\left(\left(\frac{2}{3}b + \frac{1}{12} + \frac{1}{6}q\right) + \left(\frac{3}{4} + \frac{1}{6}q\right)\right)\right) db\, ds$$

$$= \frac{3}{64}q - \frac{1}{32}q^2 - \frac{1}{48}q^3 + \frac{9}{128}$$

Due to symmetry we will also have

$$\pi_B = \frac{3}{64}q - \frac{1}{32}q^2 - \frac{1}{48}q^3 + \frac{9}{128}$$

The initial allocation which maximizes the average surplus is therefore

$$q^* = \arg\max_{q \in [0,\frac{1}{2}]} [\pi_S + \pi_B] = \frac{1}{2}$$

∎

Proof of Proposition 8:

Whenever the bidding is not extreme (i.e. $x \le b \le 1$ and $0 \le s \le 1 - x$, where x is defined by (4.7)), the trading price for the entire asset is

$$p(s,b) = \frac{r_S(s,b) + r_B(s,b)}{2} = \frac{1}{3}(s+b) + \frac{1}{6} + \frac{1}{2}\varepsilon - \frac{1}{2}\tau$$

by (4.5) and (4.6). For $q \in [0, \frac{1}{2}]$, the overall expected payoffs of each of the players is again the sum of the expected payoffs in 5 mutually exclusive and exhaustive regions:

1) When B gets the entitlement i.e. $r_B > r_S$ (this can only happen when the bids are not extreme i.e. $x \le b \le 1$ and $0 \le s \le 1 - x$ and $b - s > x$)

2) When S gets the entitlement (i.e. $r_S \ge r_B$) and

2a) Bids are not extreme i.e. $x \le b \le 1$ and $0 \le s \le 1 - x$ and $r_s \ge r_b \Leftrightarrow b - s \le x$

2b) Only the bid of B is extreme i.e. $0 \le b < x$ and $0 \le s \le 1 - x$

2c) Both bids are extreme i.e. $0 \le b < x$ and $1 - x < s \le 1$

2d) Only the bid of S is extreme i.e. $x \le b \le 1$ and $1 - x < s \le 1$

Denote by $f_B(\varepsilon, \tau)$, $f_S(\varepsilon, \tau)$ the actual ex ante payoffs (not the perceived ones!) of B and S in this equilibrium.

33

$$f_S(\varepsilon, \tau) = \int_0^{1-x} \int_{s+x}^1 ((1-q)(p(s,b) - s)\, dbds + \int_0^{1-x} \int_x^{s+x} q\,(s - p(s,b))\, dbds$$

$$+ \int_0^{1-x} \int_0^x q\left(s - \frac{1}{2}\left(\left(\frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau\right) + \left(\frac{2}{3}s + \frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau\right)\right)\right) dbds$$

$$+ \int_{1-x}^1 \int_0^x q\left(s - \frac{1}{2}\left(\left(\frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau\right) + \left(\frac{3}{4} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau\right)\right)\right) dbds$$

$$+ \int_{1-x}^1 \int_x^1 q\left(s - \frac{1}{2}\left(\left(\frac{2}{3}b + \frac{1}{12} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau\right) + \left(\frac{3}{4} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau\right)\right)\right) dbds$$

$$= -\frac{1}{384}(1-q)(\beta - 6(1 + \varepsilon - \tau))\beta^2 - \frac{1}{96}q(3\varepsilon - q)\beta^2 + \frac{1}{48}q(3\varepsilon - q)\beta(\beta - 4)$$

$$+ \frac{1}{128}q(\beta + 4(\tau - \varepsilon))(\beta - 4)^2 - \frac{1}{192}q\beta(\beta - 6(\varepsilon - \tau))(\beta - 4)$$

where $\beta(q, \varepsilon, \tau) = 2q + 3(1 - \varepsilon - \tau)$

$$f_B(\varepsilon, \tau) = \int_0^{1-x} \int_{s+x}^1 ((1-q)(b - p(s,b))\, dbds + \int_0^{1-x} \int_x^{s+x} q\,(p(s,b) - b)\, dbds$$

$$+ \int_0^{1-x} \int_0^x q\left(\frac{1}{2}\left(\left(\frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau\right) + \left(\frac{2}{3}s + \frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau\right)\right) - b\right) dbds$$

$$+ \int_{1-x}^1 \int_0^x q\left(\frac{1}{2}\left(\left(\frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau\right) + \left(\frac{3}{4} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau\right)\right) - b\right) dbds$$

$$+ \int_{1-x}^1 \int_x^1 q\left(\frac{1}{2}\left(\left(\frac{2}{3}b + \frac{1}{12} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau\right) + \left(\frac{3}{4} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau\right)\right) - b\right) dbds$$

$$= -\frac{1}{384}(1-q)(\beta - 6(1 - \varepsilon + \tau))\beta^2 + \frac{1}{96}q(q - 3\tau)\beta^2 - \frac{1}{192}q\beta(\beta + 6(\varepsilon - \tau))(\beta - 4)$$

$$+ \frac{1}{128}q(\beta + 4(\varepsilon - \tau))(\beta - 4)^2 - \frac{1}{48}q(q - 3\tau)(\beta - 4)\beta$$
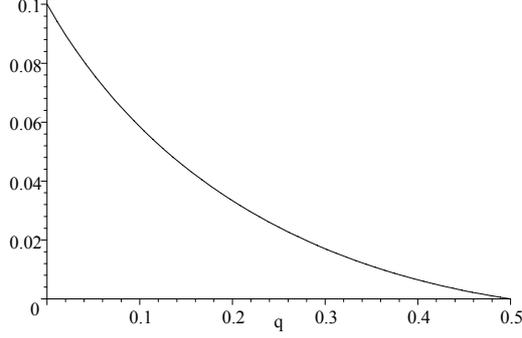
The reaction functions in the game with payoff functions f_S, f_B are

$$\tau(\varepsilon) = \arg\max_\tau f_B(\varepsilon, \tau) = \frac{14}{27}q - \frac{5}{9}\varepsilon + \frac{5}{9} - \frac{4}{27}\sqrt{(q^2 - 6\varepsilon q + 42q + 9\varepsilon^2 - 18\varepsilon + 9)}$$

$$\varepsilon(\tau) = \arg\max_\varepsilon f_S(\varepsilon, \tau) = \frac{14}{27}q - \frac{5}{9}\tau + \frac{5}{9} - \frac{4}{27}\sqrt{(q^2 - 6\tau q + 42q + 9\tau^2 - 18\tau + 9)}$$

This game is dominance-solvable, with the equilibrium biases

$$\varepsilon^* = \tau^* = \left(\frac{3}{10} + \frac{1}{3}q - \frac{1}{15}\sqrt{(9 + 80q)}\right) \tag{A.10}$$

Here is the graph of $\varepsilon^* = \tau^*$ as a function of q where $(q, 1-q)$ are the initial entitlements:

We can see that the biases $\varepsilon^* = \tau^*$ tend to zero as the initial shares $(q, 1-q)$ tend to $(\frac{1}{2}, \frac{1}{2})$. Thus in a case of equally shared property rights in the asset no biases will emerge.

Knowing the endogenous extent of the dispositions as a function of q, the surplus-maximizing initial allocations are defined by

$$
\begin{aligned}
q^* &= \arg\max_{q \in [0, \frac{1}{2}]} \left[f_S(\varepsilon^*, \tau^*) + f_B(\varepsilon^*, \tau^*) \right] \\
&= \arg\max_{q \in [0, \frac{1}{2}]} \left[\frac{1}{500} \left(80q + \left(9 - \frac{40}{3}q \right) \sqrt{80q + 9} + 27 \right) \right] = \frac{1}{2}
\end{aligned}
$$

■

Proof of Proposition 9:

The expected perceived payoffs (happiness) of S and B are

$$
\begin{aligned}
h_S(\varepsilon, \tau) &= \int_0^{1-x} \int_{s+x}^1 ((1-q)(p(s,b) - (s - \varepsilon)))\, db\, ds + \int_0^{1-x} \int_x^{s+x} q\, (s - \varepsilon - p(s,b))\, db\, ds \\
&+ \int_0^{1-x} \int_0^x q \left(s - \varepsilon - \frac{1}{2} \left(\left(\frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau \right) + \left(\frac{2}{3}s + \frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau \right) \right) \right)\, db\, ds \\
&+ \int_{1-x}^1 \int_0^x q \left(s - \varepsilon - \frac{1}{2} \left(\left(\frac{1}{4} - \frac{1}{6}q + \frac{3}{4}\varepsilon - \frac{1}{4}\tau \right) + \left(\frac{3}{4} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau \right) \right) \right)\, db\, ds \\
&+ \int_{1-x}^1 \int_x^1 q \left(s - \varepsilon - \frac{1}{2} \left(\left(\frac{2}{3}b + \frac{1}{12} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau \right) + \left(\frac{3}{4} + \frac{1}{6}q + \frac{1}{4}\varepsilon - \frac{3}{4}\tau \right) \right) \right)\, db\, ds \\
&= f_S(\varepsilon, \tau) + \frac{1}{2}\varepsilon \left(1 - 2q + x(x - 2) \right)
\end{aligned}
$$

and

35

$$h_B(\varepsilon, \tau) = \int_0^{1-x} \int_{s+x}^1 \left((1-q)(b+\tau-p(s,b))\right) dbds + \int_0^{1-x} \int_x^{s+x} q\left(p(s,b)-(b+\tau)\right) dbds$$

$$+ \int_0^{1-x} \int_0^x q\left(\frac{1}{2}\left(\left(\frac{1}{4}-\frac{1}{6}q+\frac{3}{4}\varepsilon-\frac{1}{4}\tau\right)+\left(\frac{2}{3}s+\frac{1}{4}-\frac{1}{6}q+\frac{3}{4}\varepsilon-\frac{1}{4}\tau\right)\right)-(b+\tau)\right) dbds$$

$$+ \int_{1-x}^1 \int_0^x q\left(\frac{1}{2}\left(\left(\frac{1}{4}-\frac{1}{6}q+\frac{3}{4}\varepsilon-\frac{1}{4}\tau\right)+\left(\frac{3}{4}+\frac{1}{6}q+\frac{1}{4}\varepsilon-\frac{3}{4}\tau\right)\right)-(b+\tau)\right) dbds$$

$$+ \int_{1-x}^1 \int_x^1 q\left(\frac{1}{2}\left(\left(\frac{2}{3}b+\frac{1}{12}+\frac{1}{6}q+\frac{1}{4}\varepsilon-\frac{3}{4}\tau\right)+\left(\frac{3}{4}+\frac{1}{6}q+\frac{1}{4}\varepsilon-\frac{3}{4}\tau\right)\right)-(b+\tau)\right) dbds$$

$$= f_B(\varepsilon, \tau) + \frac{1}{2}\tau\left(1-2q+x(x-2)\right)$$

respectively, where $x = \frac{1}{4} - \frac{1}{2}q + \frac{3}{4}(\varepsilon + \tau)$ as before.

As the endogenous extent of the dispositions ε^*, τ^* as a function of q is as in (A.10), the happiness-maximizing initial allocations are defined by

$$q^{**} = \arg\max_{q\in[0,\frac{1}{2}]} \left[h_S(\varepsilon^{**}, \tau^{**}) + h_B(\varepsilon^{**}, \tau^{**})\right]$$

$$= \arg\max_{q\in[0,\frac{1}{2}]} \left[\frac{1}{750}\left(-345q - 300q^2 + (18+55q)\sqrt{80q+9} + 54\right)\right] = \frac{1}{2}$$

∎

References

[1] Arlen, J., M. Spitzer and E. Talley (2002), "Endowment Effects Within Corporate Agency Relationships", *Journal of Legal Studies* **31(1)**.

[2] Ayres I. and E. Talley (1995), "Solomonic Bargaining: Dividing a Legal Entitlement to Facilitate Coasean Trade", *Yale Law Journal* **104(5)**.

[3] Babcock L. and G. Loewenstein (1997), "Explaining Bargaining Impasse: The Role of Self-Serving Biases", *Journal of Economic Perspectives* **11(1)**, pp. 109-126.

[4] Baird D.G., Jackson T.H. and Adler B. (2000), *Cases, Problems, and Materials on Bankruptcy*, Foundation Press.

[5] Bar-Gill, O. and C. Fershtman (2001), "The Limit of Public Policy: Endogenous Preferences," Foerder working paper 5-01, Tel Aviv University.

[6] Becker, G. (1976), "Altruism, egoism and genetic fitness: Economics and sociobiology," *Journal of Economic Literature* **14**, pp. 817-826.

[7] Becker, G. (1996), *Accounting for Tastes,* Harvard University Press.

[8] Benos A.V. (1998), "Aggressiveness and Survival of Overconfident Traders," *Journal of Financial Markets*, **1**, pp. 353-383.

[9] Bergman N. and Y. Bergman (2000), "Ecologies of Preferences with Envy as an Antidote to Risk-Aversion in Bargaining," mimeo, The Hebrew University of Jerusalem.

[10] Bester H. and W. Güth (1998), "Is Altruism Evolutionary Stable?" *Journal of Economic Behavior and Organization*, **34(2),** pp. 211-221.

[11] Bolle F. (2000), "Is Altruism Evolutionarily Stable? And Envy and Malevolence? - Remarks on Bester and Güth," *Journal of Economic Behavior and Organization*, **42**(1), pp. 131-133.

[12] Camerer, C., G. Loewenstein & M. Rabin (2003), *Advances in Behavioral Economics* (Princeton Press, forthcoming).

[13] Carmichael, H. L. and W.B. MacLeod (2002), "How Should a Behavioral Economist Do Welfare Economics?", Queen's University mimeo.

[14] Chatterjee K. and W. Samuelson (1983), "Bargaining under Incomplete Information", *Operations Research* **31**, pp. 835-851.

[15] Cooter, R. D. and A. Porat (2002), "Anti-Insurance", Berkeley Olin Program in Law & Economics, Working Paper Series. Working Paper 49 (available for download at: http://repositories.cdlib.org/blewp/49).

[16] Cramton, P. (1985), "Sequential Bargaining Mechanisms" in Game-Theoretic Models of Bargaining, Alvin Roth (ed.), *Cambridge University Press*, Chapter 8, pp. 149–179.

[17] Cramton, P., Gibbons, R. and P. Klemperer (1987), "Dissolving a Partnership Efficiently", *Econometrica* **55(3)**, pp. 615-632.

[18] Dawid, H. and W.B. MacLeod (2002), "Evolutionary Bargaining with Cooperative Investments", Working Paper, USC.

[19] Dekel E., J. Ely, and O. Yilankaya (1998), "Evolution of Preferences," mimeo, Northwestern University.

[20] Dollar, D. (2000), "What can Research Tell Us about Making Lending Aid More Effective?" (World Bank Working Paper, Washington DC).

[21] Ely J. and O. Yilankaya (2001), "Nash Equilibrium and the Evolution of Preferences," *Journal of Economic Theory* , **97**(2), pp. 255-272.

[22] Fershtman C. and A. Heifetz (2002), "Read My Leaps, Watch for Leaps: A Theory of Endogenous Political Instability," Foerder working paper 8-02, Tel Aviv University.

[23] Fershtman C. and Y. Weiss (1997), "Why Do We Care about what Others Think about Us?," in: Ben Ner, A. and L. Putterman (eds.), *Economics, Values and Organization*, Cambridge University Press, Cambridge MA.

[24] Fershtman C. and Y. Weiss (1998), "Social Rewards, Externalities and Stable Preferences," *Journal of Public Economics*, **70**, pp. 53-74.

[25] Forman, S. and S. Patrick, eds. (2000), *Good Intentions - Pledges of aid for Post-conflict Recovery*, Lynne Rienner Publishers: London, UK

[26] Fudenberg, D. and J. Tirole (1991), Game Theory, *MIT Press.*

[27] Güth W. and M. Yaari (1992), "Explaining Reciprocal Behavior in Simple Strategic Games: An Evolutionary Approach," in Witt, U. (ed.), *Explaining Forces and Changes: Approaches to Evolutionary Economics,* University of Michigan Press.

[28] Guttman, J.M. (2000), "On the Evolutionary Stability of Preferences for Reciprocity," *European Journal of Political Economy* **16**, pp. 31-50.

[29] Harris, C. and D. Laibson (forthcoming), "Dynamic Choices of Hyperbolic Consumers," *Econometrica*

[30] Heifetz A. and E. Segev (2001), "The Evolutionary Role of Toughness in Bargaining," Forthcoming in *Games and Economic Behavior.*

[31] Heifetz A., Shannon C. and Y. Spiegel (2003), "What to Maximize if You Must", mimeo, Tel Aviv University

[32] Horowitz, J.K. and K.E. McConnel (2002), "A Review of WTA/WTP Studies," *Journal of Environmental Economics and Management* **44**, pp. 426-447.

[33] Huck S. and J. Oechssler (1998), "The Indirect Evolutionary Approach to Explaining Fair Allocations," *Games and Economic Behavior*, **28**, pp. 13-24.

[34] Jolls C. (1998) "Behavioral Economic Analysis of Redistributive Legal Rules" *Vanderbilt Law Review* **51,** p. 1653**.**

[35] Kahneman, D., Knetch, J.L and R.H. Thaler (1990), "Experimental Tests of the Endowment Effect and the Coase Theorem", *Journal of Political Economy* **98**, pp. 1325-1348.

[36] Koçkesen L., E.A.. Ok, and R. Sethi (2000a), "Evolution of Interdependent Preferences in Aggregative Games," *Games and Economic Behavior*, **31**, pp. 303-310.

[37] Koçkesen L., E.A. Ok, and R. Sethi (2000b), "The Strategic Advantage of Negatively Interdependent Preferences," *Journal of Economic Theory,* **92**, pp. 274-299.

[38] Kreps, D. (1997), "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*, **87**(2): 359 364.

[39] Kyle A.S. and A. Wang (1997), "Speculation Duopoly with Agreement to Disagree: Can Overconfidence Survive the Market Test?" *The Journal of Finance,* **LII**, pp. 2073-2090.

[40] Lester, G. and E. Talley (2000), "Trade Secrets and Mutual Investments", USC Law School Working Paper # 00-15; Georgetown Law and Economics Research Paper No. 246406.

[41] Myerson R. and M. Satterthwaite (1983), "Efficient Mechanisms for Bilateral Trading", *Journal of Economic Theory* **29**, pp.265-281.

[42] Possajennikov A. (2000), "On The Evolutionary Stability of Altruistic and Spiteful Preferences," *Journal of Economic Behavior and Organization*, **42**(1) pp. 125-129.

[43] Rabin, M. (1998) "Psychology and Economics", J*ournal of Economic Literature*, **36**(1), 11-46.

[44] Rachlinski, J. & F. Jourden, "Remedies and the Psychology of Ownership," 52 Vanderbilt Law Review 1541-82 (1998).

[45] Rotemberg, J.J. (1994) "Human Relation in the Workplace," *Journal of Political Economy*, **102**, pp. 684-717.

[46] Samuelson, L. (2001), "Introduction to the Evolution of Preferences," *Journal of Economic Theory*, **97**, pp. 225-230.

[47] Sethi, R. and E. Somanathan (2001), "Preference Evolution and Reciprocity," *Journal of Economic Theory*, **97**, pp. 273-297.

[48] Sunstein C.R. (2002) "Switching the Default Rule," *N.Y.U. Law Review,* **77,** p. 106.

[49] Talley, E. (1998), "Turning Servile Opportunities: A Strategic Analysis of the Corporate Opportunities Doctrine", *Yale Law Journal* **108(2)**.

[50] Tversky A. and D. Kahneman (1986), "Rational Choice and the Framing of Decisions", *Journal of Business* **59**, pp. 251-278.