Check for
updates

A Theory of Just Market Exchange

Ricardo Andrés Guzmán[1] (ORCID) **· Michael C. Munger**[2,3]

Any plausibly just market exchange must balance two conflicting moral considerations: *non-worseness* (Wertheimer, 1999) and *euvoluntariness* (true voluntariness; Munger, 2011). We propose an analytical theory of just market exchange that partly resolves this conflict.

Non-worseness implies that all voluntary and mutually beneficial exchanges are morally permissible. The reason is that it cannot be immoral for two negotiators to benefit one another by exchanging, if each has the right to not exchange, and is therefore not obligated to provide a benefit to the other negotiator.

Euvoluntariness, on the other hand, elaborates the idea of voluntariness to remedy deficiencies in the standard conception. *Voluntariness* requires that neither party to a negotiation is coerced into exchange by threat of direct harm, such as physical aggression or the disclosure of embarrassing secrets.[1] *Euvoluntariness* further

---

[1] More precisely, Guzman and Munger (2014) argued that voluntariness requires five conditions to be met: the parties own the objects being exchanged; the parties have the practical and legal capacity to transfer the ownership of those objects; there is no deception or fraud, and no compulsion, addiction, or insanity; there are no large-scale uncompensated non-pecuniary externalities, and there is no coercion by threat of direct harm.

✉ Michael C. Munger
   munger@duke.edu

   Ricardo Andrés Guzmán
   rguzman@udd.cl

[1] Centro de Investigación en Complejidad Social, Facultad de Gobierno, Universidad del Desarrollo, Las Condes, Santiago, Chile

[2] Department of Political Science, Duke University, Durham, NC, USA

[3] Department of Economics, Duke University, Durham, NC, USA

 Springer

requires that the stronger party to the negotiation lacks unconscionable bargaining power; that is, the ability to extort an inordinate share of the gains from exchange.[2] The stronger party has unconscionable bargaining power if and only if two conditions are met:

- the weaker party is in dire need
- there is severe disparity between the two parties

Dire need means that the weaker party is pressed to exchange by the disastrous consequences of not exchanging, such as dying of thirst or hunger. Following convention, we will call this reversion outcome the "outside option." Dire need thus means that the weaker party's outside option is very bad, in an *absolute* sense. *Severe disparity* means that the stronger party's outside option is vastly superior to the outside option of the weaker party, in a *relative* sense.[3] The combination of dire need and severe disparity allows the stronger party to make an implicit threat to the weaker party: "Agree to my terms, or face disaster."

We claim that an extreme imbalance in bargaining power vitiates the voluntarity, and therefore the moral force, of any negotiation process. Even if an unconscionably powerful negotiator agrees to divide the gains from exchange in equal parts, it is he who largely decides the negotiation outcome. In practice, an unconscionably powerful negotiator is both a party to and the arbiter of a negotiation. This vice makes the negotiation procedurally unfair, regardless of the actual outcome, subject to the constraints we outline.

Negotiations that are procedurally unfair lead to morally dubious outcomes: If a procedurally unfair negotiation results in an agreement to exchange, the ensuing exchange will not be euvoluntary. And if a procedurally unfair negotiation results in a disagreement, an opportunity for mutual help will have been lost. Which of these outcomes should be preferred? According to proponents of the non-worseness principle, mutual help takes precedence over other moral considerations.

---

[2] The idea that unconscionable bargaining power can be an instrument of coercion can be traced back at least to Locke (2003) and Hume (1888). Hume proposed the following example: "A man, dangerously wounded, who promises a competent sum to a surgeon to cure him, wou'd certainly be bound to performance; tho' the case be not so much different from that of one, who promises a sum to a robber." Hume (1888; p. 125). In recent decades, this idea has attracted renewed interest among moral philosophers, most notably Frankfurt (1973; p. 71); Lyons (1975; 425–436); O'Neill (1985, 252–277); McGregor (1988, 23–50); Olsaretti (1998; 2004, 119–154), Snyder (2008, 389–405), Zwolinski (2009), and Munger (2011). More recently, Vrousalis (2013) connects exploitation with domination, seeking to define exploitation as the self-enriching instrumentalization of another's vulnerability. Finally, the rejection of substantial inequality in bargaining strength is a condition of the exchange situation, not a condition of the wealth positions of the participants in a broader sense. The issues discussed in regard to the diminishing marginal utility of wealth and the arguments for redistribution are summarized in Schmidtz (2000). We are considering only the narrow situation of the exchange itself.

[3] In the jargon of game theory, a negotiator's outside option is the best of his alternatives to a negotiated agreement. In a negotiation, the weaker party is the negotiator with the worse outside option, whereas the stronger party is the negotiator with the better outside option. There is some reason to believe that the evolved moral intuitions of humans are compatible with a "social contract" theory in which exchange is an integral, almost unconsciously accepted, part. See Mizzoni (2010).

Hence, voluntary exchanges must be considered just, even if they are not euvoluntary (Munger, 2011).

The conflict between euvoluntariness and non-worseness is patent in the following example.[4] In a poor country, a peasant family is starving. A wealthy American woman with total renal failure approaches the father of the family and offers to buy one of his kidneys, which happens to be a "match" to her blood type. The peasant is in dire need, because his outside option is family starvation. The wealthy woman, on the other hand, has a vastly superior outside option: Her husband—a compatible donor—has volunteered to give her one of his kidneys. The woman has unconscionable bargaining power because the peasant is in dire need and there is a severe disparity in the outside options of the two parties. Hence, the negotiation is not procedurally fair. If the kidney sale goes through, the exchange will be voluntary—there is no coercion involving human agency—but it will not be *euvoluntary*.[5] The peasant coerced by circumstance.

However, while the absence of euvoluntarity means the situation lacks the "moral pass" afforded truly voluntary exchanges, it does not therefore immediately follow that the kidney sale is morally objectionable, or that it should be legally blocked. After all, the woman has the right to take her husband's kidney, and the peasant has the right to keep his kidney and starve along with his family. How can it be morally wrong for these two people to help each other by exchanging? To each, the path of mutual help is preferable to no mutual help, and the other party is made better off by participating in the exchange. The exchange is unfair, because it is not euvoluntary, but this must be balanced against the concern for the welfare of the less well-off party, the peasant. It would be paradoxical if concern for the peasant consigns him and his family to the inferior outcome—starvation—when there is an outcome—exchange—that the peasant prefers and which objectively improves his welfare. Therefore, the non-worseness principle overrules the violation of euvoluntarity in this instance and implies that the kidney-for-money exchange is morally permissible.

Our theory of just market exchange balances euvoluntariness and non-worseness, providing moral guidance in hard cases like the one above. The theory applies to bilateral negotiations of marketable goods and services under conditions of voluntariness, but not necessarily of euvoluntariness. Our method is to postulate three plausible moral principles, which can be applied by using a game-theoretic model that we call the fictitious negotiation model.[6] If the stronger party to a negotiation

---

[4] Adapted from an example by Sandel (1998). For a more general view of "value" and its determination, see Gaus (1990).

[5] The validity of this example depends on whether a kidney qualifies as a marketable good or not (a marketable good is one whose commerce is both moral and legal). Since the marketability of kidneys is a controversial matter, we refrain from taking a position here. We just assume that kidneys are marketable for the sake of argument. One view of "morally marketable" by Brennan and Jaworski (2015) would hold that "if it can be done, it can be done for money." We are not necessarily endorsing that view here, but simply using an example in which kidneys are marketable for the sake of argument.

[6] This approach is consistent with a number of philosophical treatments of this problem, notably Smith's (1776/1976) "impartial spectator" or Nagel's (1991; p. 65) elaboration of the "impersonal standpoint". Interestingly, as Levy (2019) points out, the problem of "impartiality" relates to the requiring some outside and objective reasons. If participants are allowed to construct their own reasons, indeterminacy is likely to be the result. Yet it is not clear how to get around this problem, as Levy rightly notes, other than to have (we claim) analytically-based normative theories such as ours.

applies this model, the outcome of the negotiation will be just, even if the negotiation is procedurally unfair and leads to a non-euvoluntary exchange.

Three principles of just market exchange

We propose three principles of just market exchange:

1. No beneficence: A negotiator in a market setting is not morally obligated to give something that he owns in exchange for something that he values less than what he is giving.[7]
2. No abuse: If a negotiation is procedurally unfair, the stronger party is forbidden from using his unconscionable bargaining power, except where this prohibition conflicts with the first principle, in which case he can use his power only to avoid a loss.
3. Free-exchange: If a negotiation is procedurally fair, the parties have the right to exchange at any mutually agreed terms, and they also have the right to not exchange, as long as such rights do not conflict with the first and second principles.

The no-beneficence principle establishes a weak form of property rights: It grants the negotiator the right to be duly compensated for giving up things he owns, but it does not grant him the rights to oblige others to exchange; neither does it grant him a right to refuse to exchange (more on this later). Moreover, the no-beneficence principle is compatible with any duty of beneficence that the negotiator might have toward the other party, provided that this duty is not created, modified, or destroyed by market exchange. Some authors have argued, not implausibly, that market duties and rights are indissociable from duties of beneficence. Wells (2017), for example, argues that in emergencies the duty of beneficence can destroy the right (not) to sell help, possibly at a zero price (Wells, 2017; pp. 611-618). In other words, the no-beneficence principle isolates the duties and rights that people have as market negotiators from the duties of beneficence that they might have when they play other roles: relatives, fellow citizens, members of different social strata, and so on. We make no claim that beneficence is wrong, or that it does not exist. Rather, we address the analytically narrow setting where bargaining does not require beneficence.[8]

---

[7] It is important to note that the "no beneficence" condition is highly circumscribed to the domain of commercial exchange. One is always allowed to act on charity, and in fact there may be a *duty* of charity in extreme circumstances. Thus, we are making a distinction that is useful analytically, but is unlikely to be very appealing in actual choice situations: we are restricting our consideration to market exchange, as the title to this piece suggests. Thus, there may be a duty of beneficence or charity, but that duty arises from considerations outside the logic of market exchange. A deeper consideration of the operation of Kantian imperatives would be required for a general theory of beneficence, and that is outside our scope. For some outlines of what such a consideration would look like, see Timmermann (2005). We thank an anonymous referee for pointing out that the previous discussion of this point was inadequate and misleading.

[8] Consider the following example: A well-to-do woman has an apartment to rent and her impoverished brother has nowhere to live. He asks her to lend him the apartment while she recovers financially. The principle of non-beneficence implies that the woman can reject that request, as she would incur an opportunity cost: the forgone rent. At the same time, she could have a parallel obligation to help her brother, by given him a sum of money possibly different than the rent of the apartment. He could use that money to rent his sister's apartment or spend it on something else.

The no-abuse principle regulates the exercise of unconscionable bargaining power. We say that a negotiator uses his unconscionable bargaining power if and only if he negotiates an outcome that the weaker party would reject if the negotiation were fair. The no-abuse principle states that using unconscionable bargaining power constitutes abuse, except when its use is the only way to avoid a loss.

Prima facie, the free-exchange principle affirms a strong form of property rights, as it gives negotiators the freedom to buy or sell under any mutually agreed terms, as well as the freedom to refuse to buy or sell. However, these freedoms are subordinated to the higher duty not to abuse others. This duty, in turn, is limited by a weaker form of property rights, specified by the no-beneficence principle.

The three principles of just market exchange strike a balance between the non-worseness principle and euvoluntariness. Like the non-worseness principle, the three principles always allow the parties to exchange.

Unlike the non-worseness principle, the three principles prohibit the use of unconscionable bargaining power during negotiations. In this sense the ensuing exchanges resemble euvoluntary exchanges. There is one exemption from this prohibition: The stronger party is allowed to use his unconscionable bargaining power to avoid a loss, in which case the ensuing exchange will not resemble a euvoluntary exchange. Nevertheless, this bargaining behavior does not count as abuse, because no negotiator is morally obligated to harm himself for the benefit of another.

The fictitious negotiation model

The three principles of just market exchange can be applied by means of a fictitious negotiation model. Below we summarize this model in informal terms, postponing its mathematical formulation to the next section.

A fictitious negotiation is a mental transformation of an actual negotiation. It represents a hypothetical scenario in which the weaker party has a fictitious outside option. By maximizing the value of the fictitious outside option, the transformation minimizes the procedural unfairness of the fictitious negotiation.[9]

---

[9] The intuition behind the fictitious negotiation model can be found in a short text (under 2,000 words) scribbled by John Locke inside the cover of a notebook. The text is captioned "Venditio"--Latin for "sale"—and dated 1695 by a later auditor. It seems to be a quick set of notes for a full essay that Locke never elaborated. In Venditio, Locke analyzes four examples of hard bargains and draws conclusions about the just price. His analysis is based on a sophisticated conception of voluntariness that anticipates the notion of euvoluntariness. In one of his examples, he asks rhetorically: "A ship at sea that has an anchor to spare meets another which has lost all her anchors. What here shall be the just price that she shall sell her anchor to the distressed ship? To this I answer the same price that she would sell the same anchor to a ship that was not in that distress." In other words, Locke responds that the rescuer ship must negotiate with the distressed ship as if the distressed ship had a better outside option. For an in-depth analysis of Venditio, see Ricardo Guzman and Michael Munger, op. cit. Wertheimer proposes a similar device to determine the fair price in bilateral negotiations. He claims that the fair price corresponds to the price that would prevail in a hypothetical competitive market. According to Wertheimer, the hypothetical market price is a good benchmark for fairness because in competitive markets buyers and sellers are price takers. Therefore, neither party to a negotiation has the power to impose the terms of the exchange. See Alan Wertheimer, op. cit, p. 232.

This maximization is subject to three constraints. First, the fictitious outside option must be at least as good as the actual outside option of the weaker party. Second, the fictitious outside option must not be improved beyond the point where the stronger party loses his unconscionable bargaining power. Beyond that point the fictitious negotiation becomes procedurally fair, so further improving the fictitious outside option does not serve the purpose of reducing unfairness. Third, the fictitious outside option must not be improved beyond the point where the fictitious negotiation has zero gains from exchange. This constraint preserves the incentives to exchange in the fictitious negotiation.[10] The fictitious negotiation model gives the following instruction to the stronger party: Negotiate as if the weaker party actually had the fictitious outside option, while ignoring any considerations of justice. If the stronger party follows this instruction, he will comply with the three principles of just market exchange. Apart from the three principles, the model imposes no additional constraints on bargaining behavior. Therefore, the fictitious negotiation model is more than a method to put the three principles into practice: It is an algorithmic representation of the three principles. The mathematical formulation of the theory of just market exchange requires a formal definition for the concept of unconscionable bargaining power. We say that the stronger party has unconscionable bargaining power if and only if the difference in the (values of the) outside options exceeds a disparity threshold. The difference in the outside options captures the notion of disparity, while the disparity threshold captures the notion of need. The worse the weaker party's outside option, the greater his need, and the lower the disparity threshold. A lower disparity threshold implies that smaller differences in the outside options will suffice to create unconscionable bargaining power and make the negotiation procedurally unfair.

Parametricity and modularity

The theory of just market exchange is parametric: The free parameter is the observer's moral aversion to disparity in market negotiations. A specific type of disparity aversion is captured by a specific shape of the disparity threshold function. We take no position on what value of this moral intuition is "right," but rather offer a modular approach of an "if-then" sort: If one has a particular level of moral aversion to disparity, then certain behavioral choices are morally obliged.

Even though disparity aversion is idiosyncratic and subjective, it need not be arbitrary. In some cases, it can be deduced from the observer's moral principles. For example, radical egalitarians could make an argument from exploitation, concluding that the disparity threshold must be zero for all values of the weaker party's outside option. From this it follows that the stronger party is virtually always unconscionably powerful, that virtually all negotiations are procedurally unfair, and that virtually no market exchange is euvoluntary. Analogously, at the other extreme, radical libertarians could make an argument from self-ownership and conclude that the disparity threshold must be infinitely high for all values of the weaker party's outside option. From this it follows that the stronger party is never unconscionably powerful, that all negotiations are procedurally fair, and that all market exchanges are euvoluntary.

---

[10] It is a general property of formal negotiation models that the gains from exchange diminish and eventually turn negative as the parties' outside options improve.

The theory of just market exchange is designed to be a module that regulates a specific aspect of social interaction: market negotiations. This module can be assembled to other modules that deal with the many domains of moral or political action, such as family life, citizenship, and social hierarchy. Like the parts of a mountain bike, each moral module has different versions that are interchangeable, although not all versions are compatible. Every combination of compatible modules constitutes a particular moral or political philosophy. We advocate no one value of the parameter; our approach is to enable a generic representation of many moral intuitions in the same model.

The rest of this paper is organized as follows. In Section 1 we formalize the theory of just market exchange, and we derive the rules of conduct implied by its three principles. We also formulate the fictitious negotiation model, and we prove that it implements the three principles of just market exchange. In Section 2 we apply the model to three archetypal cases: competitive market exchange, a rescue at sea, market exchange between relatives, and sweatshop labor. Finally, in Sect. 3 we summarize the rules of conduct implied by the theory, and we discuss its policy implications.

1 The Theory

1.1 A Formal Negotiation Model and a Taxonomy of Negotiation Outcomes

A two-party negotiation is a three-tuple $\left(\gamma, d_w, d_s\right) \in \mathbb{R}^3$ where γ represents the gains from exchange to be divided between the parties, and d_w and d_s represent the parties' outside options. Subscript w identifies the weaker party, while subscript s identifies the stronger party. The stronger party has a better outside option, which means that

$$d_w < d_s \tag{1}$$

A negotiation outcome is an allocation $(x_w, x_s) \in \mathbb{R}^2$, where x_w and x_s are the parties' payoffs. The final negotiation outcome is reached through an unspecified process of voluntary offers and counteroffers called bargaining.

The benefit from exchange, or *profit*, of party i is the difference between his payoff and his outside option:

$$\pi_i = x_i - d_i, \tag{2}$$

where i can be either w or s. Profits can be positive, negative, or zero, depending on the values of x_i and d_i. If the parties fail to reach an agreement, they get their outside options, which implies that both parties make zero profit (the status quo or reversion outcome).

There are two kinds of negotiation outcomes: disagreement and agreements.

1. The *disagreement* consists of the two outside options: (d_w, d_s).
2. The set of *agreements* is given by

$$A = \left\{(x_w, x_s) : \pi_w + \pi_s = \gamma\right\} \tag{3}$$

Definition (3) says that in case of an agreement the parties' profits must add up to the gains from exchange. Using equation (2), we can restate this definition as follows:

$$A = \left\{ \left(x_w, x_s\right) : x_w + x_s = d_w + d_s + \gamma \right\} \tag{4}$$

Irrespective of the negotiation process, all negotiations eventually end in a feasible negotiation outcome.

The set of **feasible negotiation outcomes** are then the union of all possible agreements and disagreements:

$$X = A \cup \left\{ \left(d_w, d_s\right) \right\} \tag{5}$$

Combining definitions (4) and (5) we get

$$X = \left\{ \left(x_w, x_s\right) : x_w + x_s = d_w + d_s + \gamma \right\} \cup \left\{ \left(d_w, d_s\right) \right\}. \tag{6}$$

An example shows how an informal narrative illustrates a negotiation model.

Example 1 (A sculpture for sale) *William has carved a sculpture, to which he attaches a sentimental value of $1,000. The rest of his wealth (material and immaterial) amounts to $1,000,000. Samantha, a collector, is interested in the sculpture, which she values at $1,200. Samantha's wealth (material and immaterial) amounts to $1,500,000. No other buyer is interested in the statue.*

Based on this information, we will calculate the gains from exchange and the outside options. We will also identify the sets of feasible negotiation outcomes and agreements.

The gains from exchange are equal to the difference in valuations:

$$\gamma = \$1,200 - \$1,000 = \$200. \tag{7}$$

William's outside option corresponds to his wealth plus the value he attaches to the sculpture:

$$d_w = \$1,000,000 + \$1,000 = \$1,001,000. \tag{8}$$

Samantha's outside option corresponds to her wealth:

$$d_s = \$1,500,000 \tag{9}$$

Since d_w is less than d_s, William is the weaker party to the negotiation.

Using definition (4) and equations (7)-(9), we formulate the set of agreements, which contains the surplus to be assigned:

$$A = \left\{ \left(x_w, x_s\right) : x_w + x_s = \$2,501,200 \right\} \tag{10}$$

Finally, using definition (5), equations (8) and (9), and definition (10), we construct the set of feasible negotiation outcomes:

$$X = \left\{ (x_w, x_s) : x_w+, x_s = \$2,501,200 \right\} \cup \left\{ (\$1,001,000, \$1,500,000) \right\}.$$
(11)

This describes the formal aspects of an arbitrary negotiation model.

Negotiations can be classified according to their profitability:

1. If $\gamma > 0$, then the negotiation is **profitable**.
2. If $\gamma = 0$, then the negotiation has **zero-profit**.
3. If $\gamma < 0$, then the negotiation is **unprofitable**.

An essential feature of market negotiations is that they are not unprofitable. Thus, we will work under the assumption that the gains from exchange are non-negative; that is, we will assume that

$$\gamma \geq 0$$
(12)

According to the distributions of the gains from exchange, agreements fall into two types: In a **non- detrimental agreement** neither party loses; in a **detrimental agreement** at least one party loses. A non-detrimental agreement is feasible if and only if $\gamma \geq 0$. A detrimental agreement, on the other hand, is always feasible: It only requires that one party incurs a loss voluntarily for the benefit the other party. Such transactions are cases of gift-giving or supererogatory charity, not of erogatory market exchange. Therefore, we will concentrate on the non-detrimental agreements. The set of non-detrimental agreements is given by:

$$A_{ND} = \left\{ (x_w, x_s) \in A : x_w \geq d_w \wedge x_s \geq d_s \right\}$$
(13)

But definition (4) says that all agreements satisfy the following condition:

$$x_s = d_w + d_s + \gamma - x_w$$
(14)

Using this equation, we can restate definition (13) as follows:

$$A_{ND} = \left\{ (x_w, x_s) : d_w \leq x_w \leq d_w + \gamma \wedge x_s = d_w + d_s + \gamma - x_w \right\}.$$
(15)

By definition, a disagreement is non-detrimental. Thus, the set of **non-detrimental outcomes** includes all the non-detrimental agreements and the disagreement:

$$X_{ND} = \left\{ (x_w, x_s) : d_w \leq x_w \leq d_w + \gamma \wedge x_s = d_w + d_s + \gamma - x_w \right\} \cup \left\{ (d_w, d_s) \right\}.$$
(16)

Figure 1 is a graphical representation of the negotiation model. It illustrates the two relevant cases: profitable negotiations and zero-profit negotiations. In each case, the figure distinguishes the sets of detrimental and non-detrimental outcomes, and the disagreement. Note that a zero-profit negotiation has a unique non-detrimental agreement, which is equivalent to a disagreement. This equivalence follows directly from definition (16).

We will assume that all non-detrimental agreements are feasible, or potentially achievable through some bargaining process, although this is not always the case in reality. In actual bargaining settings, it may be that there are gains from exchange,

Fig. 1 Graphical representation of the negotiation model: (**a**) a profitable negotiation, (**b**) a zero-profit negotiation
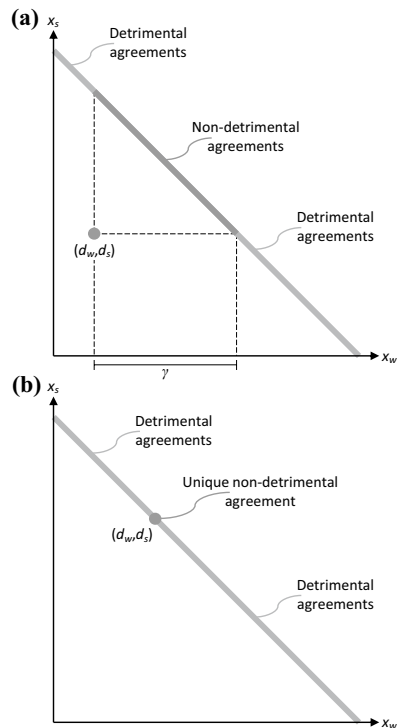
but that they cannot be divided between the parties; for instance, if it happens that one of the parties is cash poor and cannot borrow.

Example 2 (A baker meets a starving man) *Walter is penniless. Sandra is a prosperous baker. Walter enters Sandra's bakery and begs her for bread, which would save him from starvation. There are large gains from exchange at stake, since the value of a human life is much higher than the cost of producing a loaf of bread. However, there are only two possible negotiation outcomes, and both are detrimental: Sandra ignores Walter and he starves, or she gives him a loaf of bread for free (losing the opportunity cost value of the ingredients, or the price at which it could have been sold), and Walter survives. Since no non-detrimental agreement is achievable, this negotiation is not a true case of market exchange. Thus, the theory of just market exchange is silent about its moral content.*

1.2 Unconscionability, Procedural Fairness, and Euvoluntary Exchange

The stronger party has **unconscionable bargaining power** if and only if the disparity in outside options exceeds a disparity threshold; that is, if and only if

$$d_s - d_w > \theta, \tag{17}$$

Fig. 2 A plausible disparity
threshold function
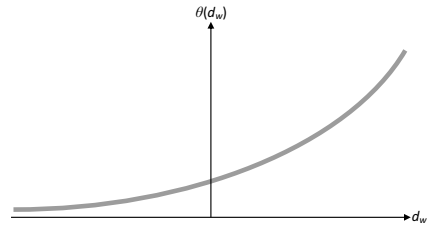
where θ is the disparity threshold.

The θ threshold is a function of the weaker party's outside option, so it must be nonnegative:

$$\theta\left(d_w\right) : \mathbb{R} \to \mathbb{R}_{\geq 0}, \tag{18}$$

Intuitively, the disparity threshold should increase with the weaker party's outside option. The reason is that the same degree of disparity creates less bargaining power for the stronger party when the weaker party has a better outside option.

Figure 2 shows a plausible disparity threshold function. As d_w grows smaller, the value of $\theta(d_w)$ approaches zero asymptotically. At the limit, even the smallest disparity in outside options will be larger than $\theta(d_w)$. This implies that virtually all negotiations will be considered unfair, and the ensuing exchanges will be considered non-euvoluntary. But for large values of d_w—meaning the outside option of the weaker party is something like normal middle class life—the value of $\theta(d_w)$ will be very high. That is, for all but the most fervent egalitarians, there will be no moral objections to normal market negotiations between parties of unequal bargaining power if the weaker party is not particularly poor or desperate. Thus, that even negotiations with significant disparity in outside options will be considered fair, and the ensuing exchanges are euvoluntary.

Exactly how much bargaining power constitutes unconscionable bargaining power depends on the observer's idiosyncratic moral aversion to disparity in outside options. This is the point of the θ parameter: it is the degree of disparity, given the particulars of the negotiation in question, where an *observer* concludes the bargain is unfair. Different types of disparity aversion are captured by different shapes of the disparity threshold function. Since this function is left unspecified, the fictitious negotiation model is compatible with the entire spectrum of disparity aversion: from radical egalitarianism to radical libertarianism.

If the stronger party has unconscionable bargaining power, the negotiation is vitiated by his ability to negotiate an inordinate share of the gains from exchange. This vice makes the negotiation procedurally unfair. Conversely, if the stronger party does not have unconscionable bargaining power, the negotiation is procedurally fair. More precisely,

1. If $d_s - d_w \leq \theta(d_w)$ the negotiation is **fair** in a procedural sense (if the inequality is strict, then the negotiation is strictly fair).
2. If $d_s - d_w = \theta(d_w)$ the negotiation is **barely fair** in a procedural sense.
3. If $d_s - d_w > \theta(d_w)$ the negotiation is **unfair** in a procedural sense.

Hereinafter, where we speak of fairness, we will be speaking of procedural fairness, in the sense defined above.

We distinguish between a negotiation, which is the process of reaching an agreement or a disagreement, and an **exchange**, which implements an agreement. If a negotiation is fair, any ensuing exchange will be **euvoluntary**. If a negotiation is unfair, any ensuing exchange will be **non-euvoluntary**.

Example 3 (Radical egalitarians and radical libertarians) *Radical egalitarianism and radical libertarianism can be depicted as limiting cases of the disparity threshold function.*

Radical egalitarians are extremely disparity averse. To them, any disparity is unfair, no matter how small it may be. Implicitly, radical egalitarians maintain that $\theta(d_w) = 0$ for all values of d_w. Therefore, they consider that virtually no voluntary market exchange is euvoluntary.

Radical libertarians lie on the opposite extreme of the disparity aversion spectrum. They reject the concept of coercion by means of unconscionable bargaining power. Implicitly, radical libertarians maintain that $\theta(dw)$ is "arbitrarily high" for all values of d_w. Therefore, they consider that any voluntary market exchange is also by definition euvoluntary.

The fairness of a negotiation can also be judged in terms of the quality of the weaker party's outside option. The benchmark is the minimum fair outside option, which is the outside option that would make the negotiation barely fair. Let $\underline{d_w}$ denote the minimum fair outside option. This variable is the solution to the following equation:

$$d_s - \underline{d_w} = \theta\left(\underline{d_w}\right), \tag{19}$$

In words, this means that the minimum fair outside option is exactly the outside option that is "barely fair." More generally:

1. If $d_w \geq \underline{d_w}$ the negotiation is fair,
2. If $d_w = \underline{d_w}$ the negotiation is barely fair,
3. If $d_w < \underline{d_w}$ the negotiation is unfair.

Not all unfair negotiations are *equally* unfair. The stronger party's ability to exert pressure on the weaker party depends on the magnitude of the gap between the weaker party's outside option and the minimum fair outside option. Thus, it makes sense to define the **degree of unfairness** of a negotiation in the following manner:

$$\delta = \max\left\{\underline{d_w} - d_w, 0\right\}. \tag{20}$$

If $\delta = 0$, the negotiation is fair. If $\delta > 0$, the negotiation is unfair.

Suppose a negotiation is unfair. The stronger party uses his unconscionable bargaining power if and only if he negotiates an outcome that the weaker party would

reject if the negotiation was fair. Another way of saying this is the following: The stronger party uses his unconscionable bargaining power if and only the negotiation is unfair ($\delta < 0$), *and* the weaker party gets less than the minimum fair outside option ($x_w < \underline{d_w}$).

1.3 Justice in Market Exchange

In this section we translate the three principles of just market exchange into rules of conduct for the stronger party. To this end, we will exploit the moral correspondence between bargaining behaviors and negotiation outcomes: A just bargaining behavior leads to a just negotiation outcome; a just negotiation outcome reveals that the negotiators behaved justly. So instead of inferring the rules of conduct directly from the three principles, we will use the three principles to identify the just negotiation outcomes. Having identified these outcomes, we will derive the rules of conduct that lead to them. These rules will instantiate the following general rule: make offers, counteroffers, rejections, and acceptances within the set of just negotiation outcomes.

Below we present the three principles of just market exchange, and the logical conditions they impose on the set of just outcomes.

1. **No beneficence:** A negotiator is not morally obligated to give something that he owns in exchange for something that he values less than what he is giving.

   All non-detrimental outcomes meet the no-beneficence principle. These outcomes satisfy the following condition:

$$d_w \leq x_w \wedge d_s \leq x_s. \tag{21}$$

   Detrimental outcomes do not meet the no-beneficence principle, because they can only be achieved if one of the two parties consents to take harm for the benefit of the other; this exceeds the demands of justice in a commercial setting.

2. **No abuse:** If a negotiation is unfair, the stronger party is forbidden from using his unconscionable bargaining power, except where this prohibition conflicts with the first principle, in which case he can use his power only to avoid a loss.

   Suppose a negotiation is unfair. In terms of outcomes, the no-abuse principle says that if the weaker party gets less than the minimum fair outside option, then the stronger party gives all the gains from exchange to the weaker party

$$\delta > 0 \rightarrow (x_w < \underline{d_w} \rightarrow x_w = d_w + \gamma) \tag{22}$$

3. Free-exchange: If a negotiation is fair, the parties have the right to exchange at any mutually agreed terms, and they also have the right to not exchange, as long as such rights do not conflict with the first and second principles.

   This principle states that a feasible negotiation outcome is just if and only if it satisfies conditions (21) and (22).

From the above principles, we can derive the set of just negotiation outcomes.

Proposition 1 *The set just negotiation outcomes is given by*

$$X_J = \{(x_w, x_s) : \min\{\max\{\underline{d_w}, d_w\}, d_w\} + \gamma\} \leq x_w \leq d_w + \gamma \wedge x_s = \{d_w + d_s + \gamma - x_w\} \cup \left\{(d_w, d_s) : \delta = 0 \vee \gamma = 0\right\}$$
(23)

Proof See Appendix A.1, available online at http://www.michaelmunger.com/paper s/TJME-Apps.pdf. □

We are also interested in the set of **just agreements**, as this set defines the moral limits of the terms of exchange. The set of just agreements is simply the set of agreements that are just:

$$A_J = A \cap X_J.$$
(24)

From definitions (4), (23), and (24) it follows that

$$A_J = \{(x_w, x_s) : \min\{\max\left\{\underline{d_w}, d_w\right\}, d_w + \gamma\} \leq x_w \leq d_w + \gamma \wedge x_s = d_w + d_s + \gamma - x_w\}$$
(25)

There are four possible **negotiation scenarios**: (1) the negotiation is fair; (2) the negotiation is unfair but profitable, and the degree of unfairness is less than the gains from exchange; (3) the negotiation is unfair but profitable, and the degree of unfairness is greater than or equal to the gains from exchange; and (4) the negotiation is unfair and has zero profit. The following proposition characterizes the set of just outcomes in each of these negotiation scenarios.

Proposition 2 *The sets of just outcomes and just agreements have the following properties:*

1. *If the negotiation is fair, then (a) the just outcomes are the same as the non-detrimental outcomes, and (b) the just agreements are the same as the non-detrimental agreements:*

$$\delta = 0 \rightarrow X_J = X_{ND} \wedge A_J = A_{ND}$$
(26)

2. *If the negotiation is unfair but profitable, and the degree of unfairness is less than the gains from exchange, then (a) the just agreements are a proper subset of the non-detrimental agreements, (b) this subset has a continuum of elements, (c) in all just agreements the weaker party gets at least the minimum fair outside option, and (d) a disagreement is unjust:*

$$0 < \delta \leq \gamma \rightarrow A_J = \left\{(x_w, x_s) \in A_{ND} : x_w \geq \underline{d_w}\right\} \subset A_{ND} \wedge |A_{ND}| = c \wedge (d_w, d_s) \notin X_J$$

3. *If the negotiation is unfair but profitable, and the degree of unfairness is greater than or equal to the gains from exchange, then (a) there is a unique just outcome, (b) this outcome corresponds to the unique just agreement, (c) in this agreement the weaker party gets all the gains from exchange, and (d) a disagreement is unjust:*

$$0 < \gamma \leq \delta \rightarrow X_J = A_J = \left\{(d_w + \gamma, d_s)\right\} \wedge (d_w, d_s) \notin X_J$$

Fig. 3 A fair negotiation: The just negotiation outcomes are the same as the non-detrimental outcomes, and the just agreements are the same as the non-detrimental agreements. The fictitious disagreement outcome is equivalent to the actual disagreement outcome
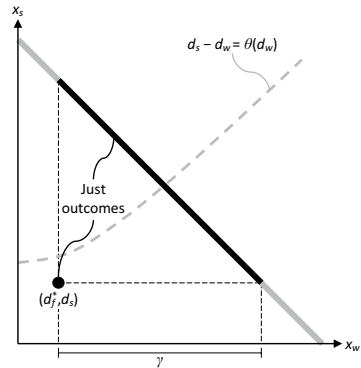

Fig. 4 An unfair but profitable negotiation, with gains from exchange that are greater than the degree of unfairness: The just agreements are a proper subset of the non-detrimental agreements, in all just agreements the weaker party gets at least the minimum fair outside option, and the disagreement outcome is unjust. The fictitious disagreement outcome is unfeasible
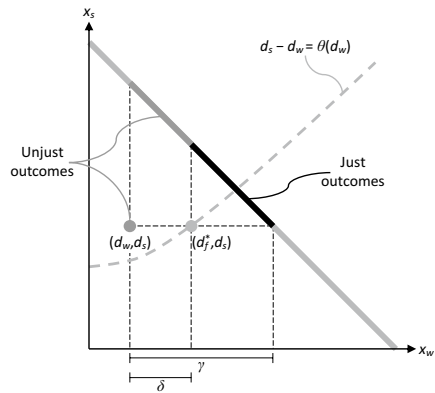
4. *If the negotiation is unfair and has zero profit, then (a) the unique non-detrimental agreement is just, and (b) a disagreement is just:*

$$\gamma = 0 \rightarrow X_J = A_J = \left\{ \left(d_w, d_s \right) \right\}$$

Proof Straightforward from definitions (15), (16), (23), and (25), and equation (20). □

Figures 3, 4, 5 and 6 illustrate this proposition (these figures also show the fictitious outside options, which will be introduced in the next section).

As a corollary, we obtain five rules of conduct that regulate bargaining behavior in voluntary market exchanges.

Corollary 1 *The three principles of just market exchange imply the following rules of conduct:*

1. *If the negotiation is fair, the parties have the moral right to negotiate any non-detrimental outcome.*

Fig. 5 An unfair but profitable negotiation, in which the gains from exchange are less than the degree of unfairness: There exists a unique just outcome, which corresponds to the unique just agreement, and in this agreement the weaker party gets all the gains from exchange. The just negotiation outcome is equal to the fictitious disagreement outcome, which is feasible
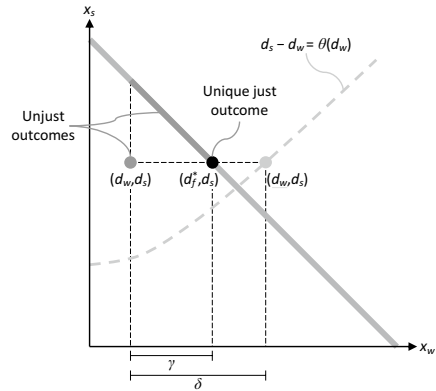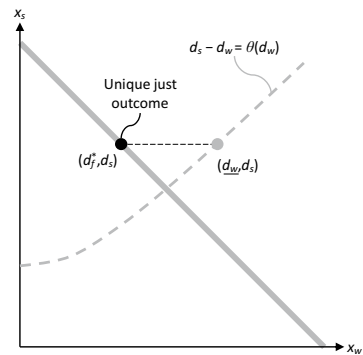
Fig. 6 An unfair negotiation with zero profit: Both outcomes are just. The fictitious disagreement outcome is equivalent to the actual disagreement outcome

2. *If the negotiation is unfair but profitable, and the degree of unfairness is less than the gains from exchange, then (a) the stronger party is morally obligated to exchange, and (b) the stronger party must give the weaker party at least the minimum fair outside option.*

3. *If the negotiation is unfair but profitable, and the degree of unfairness is greater than or equal to the gains from exchange, then (a) the stronger party is morally obligated to exchange, and (b) the stronger party must give all the gains from exchange to the weaker party.*

4. *If the negotiation is unfair and has zero profit, exchange is morally optional and the only just agreement is identical to a disagreement.*

Proof Straightforward from Proposition 2. □

Note that these rules of conduct allow market exchange in the four possible negotiation scenarios. Thus, an opportunity presented by a potential mutually beneficial exchange is never lost in the name of justice. On the contrary, if the weaker party

is unfairly disadvantaged, the stronger party cannot frustrate the opportunity of a mutually beneficial exchange. He is morally obligated to exchange, if the weaker party so requests.

1.4 How to Devise a Fictitious Negotiation

A fictitious negotiation is a mental transformation of an actual negotiation. It represents a hypothetical scenario in which the weaker party has a fictitious outside option. By maximizing the value of the fictitious outside option, the transformation minimizes the procedural unfairness of the fictitious negotiation.

More formally, the fictitious negotiation is a three-tuple $\left(\gamma_f, d_f, d_s\right) \in \mathbb{R}^3$, where γ_f are the fictitious gains from exchange, and d_f is the fictitious outside option of the weaker party. The fictitious gains from exchange and the fictitious outside option are functions of the parameters of the actual negotiation. The stronger party's outside option is the same in the fictitious negotiation as in the actual negotiation.

Proposition 3 *The fictitious gains from exchange are given by*

$$\gamma_f = \gamma + d_w - d_f \tag{27}$$

Proof See Appendix A.2, http://www.michaelmunger.com/papers/TJME-Apps.pdf. □

Note that the fictitious gains from exchange *diminish* as the fictitious outside option *improves*.

Recall that Equation (19) says that the minimum fair outside option of a negotiation only depends on the stronger party's outside option. Since the stronger party's outside is the same in the fictitious negotiation as in the actual negotiation, the minimum fair outside option is also the same in both negotiations. It follows that the fictitious degree of unfairness is given by

$$\delta_f = max\left(\underline{d_w} - d_f, 0\right)$$

We want to minimize this quantity by selecting the right value for d_f.

In formal terms, the fictitious outside option d_f solves the following **unfairness minimization problem**: Choose d to minimize

$$\max(\underline{d_w} - d_f, 0)$$

Subject to (1) $d \geq d_w$, (2) $d > d_w \rightarrow d \leq \underline{d_w}$, and (3) $\delta + d_w - d \geq 0$ (28)

The objective function is the degree of unfairness of the fictitious negotiation. The first constraint says that the fictitious outside option must be at least as good as the actual outside option of the weaker party. The second constraint says that the fictitious outside option must not be improved beyond the point where the fictitious negotiation becomes barely fair. The third constraint says that the fictitious outside

option must not be improved beyond the point where the fictitious negotiation has zero gains from exchange.

Proposition 4 The fictitious outside option is given by

$$d_f = \min\left\{\max\left\{\underline{d_w}, d_w\right\}, d_w + \gamma\right\} \tag{29}$$

Proof See Appendix A.3. http://www.michaelmunger.com/papers/TJME-Apps.pdf.
□

Again, Figures 3, 4, 5 and 6 illustrate this proposition.

The fictitious negotiation model gives the following instruction to the stronger party: Negotiate *as if* the weaker party actually had the fictitious outside option. This means that any further considerations of justice can be ignored. Apart from the three principles, the model imposes no additional constraints on bargaining behavior. Therefore, the fictitious negotiation model is more than a method to put the three principles into practice: It is an algorithmic representation of the three principles.

The proof of these claims consists in showing that the just outcomes of the actual negotiation are the same as the feasible non-detrimental outcomes of the fictitious negotiation.

Corollary 2 *The fictitious negotiation model implements the theory of just market exchange:*

$$X_J = X_{NDf} \cap X = \left\{\left(x_w, x_s\right) \,:\, d_f \leq x_w \leq d_w + \gamma \wedge x_s = d_w + \gamma - x_w\right\} \cup \left\{\left(d_w, d_s\right) \,:\, d_w = d_f^*\right\} \tag{30}$$

where X_{NDf} *is the set of fictitious non-detrimental outcomes, and* d_f^* *is the optimal value of d from equation* (28).

Proof Straightforward from definitions (16) and (23), and equation (29). □

2 Examples

2.1 A Rescue at Sea: Adapted from Locke's "Venditio"

The following example is loosely based on an adaptation of a thought experiment proposed by John Locke (2003; pp. 188–192).[11] Suppose two galleons, Wrington and Samaritan, come across each other at high seas. Wrington has lost all her

---

[11] Wertheimer (1999; pp. 232–233), discusses a similar case from real life. Two ships, the Port Caledonia and the Anna, were sheltering from a storm in Holyhead Harbour. The master of the Port Caledonia asked the Anna for a tug. After some tussle the masters agreed on a £1,000 fee, which was an unusually high price. Later, a court modified the contract, on grounds of it being extortionate. The court awarded the Port Caledonia £200 for its services.

anchors during a storm, while Samaritan has an anchor to spare. Wrington asks Samaritan to sell her an anchor. What is the maximum just price? That is, what is the highest price a person guided by moral principle in a market context could charge, recalling the caveat that the conditions proposed by Wells (2017) might imply market exchange is not the right domain of morality in some cases. But we are proposing to restrict consideration to exchange morality. We use the fictitious negotiation model developed in the previous section to propose an answer to this question. We will compare this answer with the answers offered by two other models: Wertheimer's (1999) hypothetical market price, and Locke's non-discriminatory price. Finally, we will argue that the fictitious negotiation model offers the most plausible answer.

To calculate the maximum just price for Samaritan's spare anchor, we have to consider three facts. First, in heavy seas, a galleon without an anchor is in grave danger. The galleon can easily be rolled by breaking waves; or, as she approaches the coast, the waves can smash her against the rocks. Second, having an anchor significantly reduces the probability of sinking. Third, a spare anchor is not essential for survival, but it provides valuable insurance against the accidental loss of the main anchor.[12]

These facts can be formalized as follows. Let s_k be the probability that a galleon that carries k anchors will sink during her voyage. The probability of sinking on a reef or running aground and breaking apart decreases as the galleon acquires more anchors, but the biggest improvement in safety occurs with the first anchor. Let us assume these facts are symmetric for the two ships, and the ships, cargo, and crew have equal market value in port. Then we make a comparison:

$$s_0 - s_1 > s_1 - s_2. \tag{31}$$

The left hand side is the reduction in risk for Wrington buying the first anchor; the right is the increased risk Samaritan suffers if she sold one of her anchors.

At port, the total value of a galleon includes the market price of the watercraft and her cargo, plus the monetary value of the lives of her passengers and crew. The total value of a sunken galleon is zero. At high seas, by contrast, a galleon is worth her expected value. This is equal to the galleon's total value multiplied by the probability that she successfully completes her voyage. For simplicity, we will assume that both galleons are equally valuable. The total "safe in port" value of each galleon is $z > 0$.

The galleons' outside options correspond to their expected values:

$$d_w = \left(1 - s_0\right)z, \tag{32}$$

$$d_s = \left(1 - s_2\right)z, \tag{33}$$

Since $s_O > s_1$, we have that $d_w < d_s$. This means that Wrington is the weaker party to the negotiation.[13]

Imagine that Samaritan agrees to sell her spare anchor to Wrington. When it is exchanged, Samaritan's spare anchor becomes Wrington's first or main anchor. After the exchange, both galleons are in possession of one anchor, and hence they face the same probability of sinking. By acquiring a main anchor, Wrington reduces the gold in her coffers by an amount equal to the agreed sale price. At the same time, Wrington increases her total value in the market price of the anchor, because at the end of her voyage she can sell it at the port market (or, what is the same, Wrington could "sell" the anchor to herself at the market price to use it again on a future voyage). Meanwhile, by selling her spare anchor, Samaritan increases the gold in her coffers by an amount equal to the agreed sale price. At the same time, Samaritan decreases her total value in the market price of an anchor, as she loses the opportunity to sell her spare anchor at the port market.

The galleons' allocations correspond to their new expected values:

$$x_w = (1 - s_1)(z - p + p^*) \tag{34}$$

$$x_s = (1 - s_1)(z + p - p^*) \tag{35}$$

where p is the agreed sale price, and $p^* > 0$ is the competitive market price of an anchor. Using equations (2) and (32)-(35) we calculate the (expected) profits of each galleon:

$$\pi_w = (1 - s_1)(p^* - p) + (s_0 - s_1)z, \tag{36}$$

$$\pi_s = (1 - s_1)(p - p^*) - (s_1 - s_2)z, \tag{37}$$

Adding up these profits, we obtain the gains from exchange:

$$\gamma = \left[(s_0 - s_1) - (s_1 - s_2) \right] z. \tag{38}$$

From inequality (31) it follows that the gains from exchange are positive.

Wrington's reservation price is the highest price that she can pay for an anchor without incurring a loss.

From equation (36) it follows that this price is

$$p_w = p^* + z \frac{s_0 - s_1}{1 - s_1} \tag{39}$$

Analogously, Samaritan's reservation price is the lowest price that she can charge for its spare anchor without incurring a loss. From equation (37) it follows that this price is

---

[13] One intuits that Wrington is the weaker party to the negotiation, as she is in grave danger, but this intuition can be wrong. If the cargo of Wrington was much more valuable than the cargo of Samaritan, or if Wrington carried a great number of lives while Samaritan carried a few, then Wrington would be the stronger party. Our assumption that the two galleons are equally valuable guarantees that Wrington is the weaker party.

$$p_s = p^* + z\frac{s_1 - s_2}{1 - s_2} \tag{40}$$

Note that these reservation prices have two components: the first component is the market price of an anchor; the second component is its use value for the galleon. In the case of Wrington, the use value of the anchor is the value of having a main anchor available. In the case of Samaritan, the use value of the anchor is the value of having a spare anchor available. Since a main anchor is more useful than a spare, Wrington assigns a higher use value to the anchor than Samaritan does. Therefore, Wrington's reservation price should be higher than Samaritan's reservation price.[14]

Up to this point, we have described nothing more than the material circumstances of the negotiation. But the goal is to derive the range of just prices given these material circumstances. The minimum just price follows directly from the no-beneficence principle: It is equal to Samaritan's valuation of her spare anchor, as it is unfair to require a ship to sell an anchor at a loss. Locke is very clear about this, in his statement of the problem in "Venditio":

> What here shall be the just price that she shall sell her anchor to the distressed ship? To this I answer the same price that she would sell the same anchor to a ship that was not in that distress. For that still is the market rate for which one would part with anything to anybody who was not in distress and absolute want of it. And in this case the master of the vessel must make his estimate by the length of his voyage, the season and seas he sails in, and so what risk he shall run himself by parting with his anchor, *which all put together perhaps he would not part with it at any rate*, but if he would, he must then take no more for it from a ship in distress than he would from any other. (emphasis added; from Anomaly, et al. 2016, p. 189)

To calculate the maximum just price we need an Archimedean point that informs us what the party's moral obligations are. But that is precisely what all the tiresome machinery developed in the first part of this paper was designed to provide: the moral benchmark is precisely the observer's disparity threshold function. As an example, assume that this function has the following form:

$$\theta(d_w) = \frac{1 - \beta}{\beta}\max\{d_w, 0\} \tag{41}$$

where $\beta \in (0, 1)$ is a parameter that captures the observer's disparity aversion. The more pronounced his disparity aversion, the greater the value of β.

The existence of a minimum fair outside option implies the existence of a maximum fair price. This is the maximum price such that $x_w \geq \underline{d_w}$. Using equations (19), (33), and (41), we calculate the minimum fair outside option:

$$\underline{d_w} = \beta(1 - s_2)z \tag{42}$$

---

[14] If we combine the inequality (31) and (39)–(40), it must be true that $p^* < p_s < p_w$, which confirms the claim.

Let $\bar{p}(\beta)$ denote the maximum fair price. From equations (34) and (42) it follows that

$$\bar{p}(\beta) = p^* + z - \frac{\beta(1 - s_2)z}{1 - s_1} \tag{43}$$

Note that $\bar{p}(\beta)$ is a strictly decreasing function of the observer's disparity aversion.

Analogously, the existence of a fictitious outside option that may be better than Wrington's actual outside option implies the existence of a maximum just price. This is the maximum price such that $x_w \geq d_f$. Let p_J denote the maximum just price. From equations (29) and (34) it follows that

$$p_j = p^* + z - \frac{\min\left\{\max\left(\underline{d_w}, d_w\right), d_w + \gamma\right\}}{1 - s_1} \tag{44}$$

Combining the results of equations (31) through (43) and rearranging, we can now state the maximum just price concisely:

$$p_j(\beta) = \{\max\{\min(\bar{p}(\beta), p_w\}, p_s\} \tag{45}$$

This function is decreasing in β because $\bar{p}(\beta)$ is decreasing in β.

We conclude the following:

1. The maximum just price for Samaritan's spare anchor is a decreasing function of the observer's disparity aversion.
2. Observers with low disparity aversion consider that the maximum just price is Wrington's reservation price. What is the same, Samaritan is not morally obligated to share the gains from exchange with Wrington. Observers with low disparity aversion also consider that Samaritan is morally allowed to refuse to sell her spare anchor to Wrington.
3. Observers with intermediate disparity aversion consider that the maximum just price is lower than Wrington's reservation price, but higher than Samaritan's reservation price. What is the same, Samaritan is morally obligated to share the gains from exchange with Wrington. Observers with intermediate disparity aversion also consider that Samaritan is morally obligated to sell her spare anchor to Wring- ton.
4. Observers with high disparity aversion consider that the unique just price corresponds to Samaritan's reservation price. What is the same, Samaritan is morally obligated to give all the gains from exchange to Wrington. Observers with high disparity aversion also consider that Samaritan is morally obligated to sell her spare anchor to Wrington.

The rescue at sea thought experiment can be analyzed using two other models: Wertheimer's hypothetical market price, and Locke's non-discriminatory price.

Wertheimer (1999) points out that many (perhaps most) market negotiations take place in conditions of imperfect competition, where some economic agents can use their market power to exploit others. Such is the case of Locke's rescue at sea, as

there is a single galleon without an anchor, and a single galleon with an anchor to spare. Wertheimer also points out that in competitive markets economic agents are interchangeable. Hence, they lack the capacity to exploit one another. Based on these observations, Wertheimer proposes that the non-exploitative price for a good is the price that it would have in a hypothetical competitive market (although he warns that the non-exploitative price may not be just, all things considered).

But Wertheimer overlooks two facts that compel us to conclude that his model is implausible, and that the just price could be higher than the hypothetical market price.

First, he overlooks the fact that in bilateral monopolies the seller's reservation price is equal to the competitive market price of the good plus its use value for the seller. Hence, asking Samaritan to sell her spare anchor at its competitive market price is the same as asking her to take a loss. This is an unreasonable demand in market negotiations, though of course our bright-line distinction between commerce and charity may itself be too strong. (See Wells, 2017).

The second fact that Wertheimer overlooks is that a profitable negotiation always take place under conditions of imperfect competition. Moreover, strictly speaking, all profitable negotiations are bilateral monopolies: The seller offers a good that the buyer values more than any substitute available, while the buyer is willing to pay more for this good than any other prospective buyer. So when Wertheimer claims that there is a single non-exploitative price, he tacitly claims that haggling is morally prohibited precisely in those negotiations where haggling serves a purpose: the negotiations in which there are gains to divide between the parties.

Unlike Wertheimer's hypothetical price model, the fictitious negotiation model allows haggling in many bilateral monopolies. It replaces the absolute prohibition to haggle with a less stringent constraint: If the negotiation is unfair, the model will sometimes set an upper or a lower bound to prices, leaving space for haggling between these bounds. Only in extremely severe cases of disparity and dire need will the cap be equal to the seller's or the buyer's reservation price (depending on who is the stronger party).

John Locke proposes yet another way to compute the just price for an anchor. In his original formulation of the rescue at sea thought experiment, he derives the just price using an argument of non-discrimination. He says:

> And here we see, the price which the anchor cost him, which is the market price at another place, makes no part of the measure of the price which he fairly sells it for at sea. And therefore I put in "'the place where the thing is sold'": i.e. the measure of rating anything in selling is the market price where the thing is sold. Whereby it is evident that a thing may be lawfully sold for 10, 20, nay cent per cent, and ten times more in one place than is the market price in another place perhaps not far off. (Anomaly, et al., 2016, p. 189)

Locke thus rests his entire moral case on the claim that the vessel must not discriminate against the ship in distress by asking a higher price than he would ask from an undistressed ship. In other passages, he claims that sellers are not morally obligated to sell below cost. In other words, he claims that a seller is not morally obligated to harm himself for the benefit of the buyer. Locke also makes the

empirical claim that the non-discriminatory price corresponds to the use value of the anchor for the vessel. As the vessel's reservation price is greater than the use value of the anchor, Locke's conception of just price contradicts the no-beneficence principle that he advocates.

Locke's line of argument has a questionable implication. If the distressed ship was not in distress, that is, if she had a main anchor but not a spare anchor, the vessel and the ship in distress would assign the same value to the anchor. There would be no gains from exchange. So the master of the vessel could reason as follows: "Suppose the ship in distress was not in distress. Then she values the anchor just as much as we do. Therefore, we have no moral obligation to sell her the anchor." This reasoning seems more like an excuse for not helping the ship in distress, than a mechanism to calculate the just price for the anchor.

The fictitious negotiation model leads to the opposite conclusion: If a negotiation is unfair and there are positive gains from exchange, the stronger party is morally obligated to exchange. The vessel cannot just sail away, leaving the anchorless ship to her fate.

When taken to its logical conclusion, the non-discriminatory price model entails the same implausible prescription that the non-exploitative price model: Thou shall not haggle. First, Locke claims that the minimum just price for the anchor must be the vessel's valuation of the anchor. Then he adds that "[the vessel must] take no more from [the anchor] from a ship in distress as he would take from any other." But it is possible to conceive an undistressed ship that values the anchor as much as the vessel does. This ship will not pay more for the anchor than the vessel's valuation of the anchor. Thus, the vessel's valuation of the anchor is also the maximum just price. Since the vessel's valuation of the anchor is both the minimum and the maximum just price, there exists a unique just price.

Having gone through the derivation of the maximum just price in one instance, we now present two much shorter, purely verbal examples as further illustrations.

2.2 Competitive Market Exchange Between Relatives

Wilma and Samuel are siblings. She is going through harsh economic times, while her brother is wealthy. Wilma is shopping for a used car, and it happens that Samuel wants to sell his car and buy a new one. What would be a just price for Samuel's car?

As there exists a competitive market for used cars, and in competitive markets there are zero gains from exchange, then there is a unique just price for Samuel's car: its market price (see Section 2.1). Even though his sister is dire need, Samuel is not morally obliged to give her a discount on the car.

However, this does not imply that Samuel is exempt from helping Wilma. The theory of just market exchange does not specify all the obligations between two persons: It is just one of many interchangeable modules that can be used to build up a broad moral philosophy covering different domains of human action. It is possible that one among these many modules commands Samuel to help his sister; for example, in his role as brother.

Suppose the market price of Samuel's car is $2.000. Also suppose that the module that regulates family relations obliges Samuel to give Wilma the equivalent of $1.500 in aid. Samuel could fulfill this moral obligation by writing a $1,500 check to his sister. But let us imagine that this particular module prohibits cash transfers between relatives if other ways of helping are possible (this prohibition could be based on the notion that cash transfers humiliate the recipients). Given this constraint, Samuel could take advantage of the car sale to make an implicit transfer to her sister: He could sell his car to Wilma at $500, which is $1,500 less than its market price. But Samuel has other alternatives to help his sister that do not involve selling his car at a discount: He could buy her clothes, invite her to live in his house, include her in his health insurance policy, and, in general, make gifts to his sister up to the sum of $1,500.

2.3 Sweatshop Labor

Interaction theories state that special duties of beneficence are created by economic interaction. For example, some authors have suggested that entrepreneurs have special duties toward their workers, such as paying living wages that exceed the market wage.[15] These theories are incompatible with the theory of just market exchange, which implies that the market wage is always just, provided that the labor market is competitive. This implication holds even in the case of sweatshops, which employ destitute people who are forced by circumstances to accept subsistence wages.

The main argument against interaction theories is that entrepreneurs have the right to not hire poor people, and thus to not create new, costly duties for themselves. To evade these costs, an entrepreneur can hire few skilled middle-class workers instead of many unskilled poor workers, use robots instead of human workers, or even invest his capital in another sector, such as the financial market. The result of this evasive tactic is that the poor stay unemployed, while the entrepreneur incurs no material or moral cost for taking his business elsewhere. Since interaction theories entail this absurd result, they must be revised or rejected (Wertheimer, 1999; Zwolinski, 2008).

Other theories state that people of equal wealth have the same general duties toward the poor, irrespective of their economic interaction with them.[16] Since general duties of beneficence do not interact with market duties and rights, they are compatible with our theory of just market exchange.

---

[15] Snyder (2013) and Sample (2003) advocate this idea.

[16] If the wealthy fail to fulfil those general duties, society has the option to force them by means of progressive taxation and redistribution. Fried (1981; p. 106). An interesting extension of this problem has to do with proximity: Are consumers (for example) obligated to be concerned with the condition of workers in faraway countries? That is, the fact that conditions might be a "sweatshop" if placed in one's own country may or may not imply the same obligation to be concerned if the country is distant. Weymark (2014) shows that this intuition is at best arbitrary, and may be entirely impossible to justify, at least using only distance.

3 Conclusion

We have demonstrated that the three principles of just market exchange suggest four rules of moral conduct:

1. If a negotiation is fair, the parties have the right to negotiate any non-detrimental outcome.
2. If a negotiation is unfair but profitable, and the degree of unfairness is less than the gains from exchange, then the stronger party is morally obligated to exchange, and the stronger party must give the weaker party at least the minimum fair outside option.[17]
3. If a negotiation is unfair but profitable, and the degree of unfairness is greater than or equal to the gains from exchange, then the stronger party is morally obligated to exchange, and the stronger party must give all the gains from exchange to the weaker party.
4. If a negotiation is unfair and has zero profit, exchange is morally optional and the only just agreement is identical to a disagreement.

These rules of conduct depend on a hidden parameter: the disparity threshold function, which varies from observer to observer. At first glance, this subjective element seems to allow for full arbitrariness: It seems that by choosing the right shape for the disparity threshold function you can obtain rules that allow or prohibit any conceivable bargaining behavior.

But our objective is the opposite: By embedding the disparity threshold function within an analytical framework, our theory tightly constraints the set of valid rules of conduct. Any residual subjectivity stems from fundamental philosophical discrepancies (such as the ones that separate egalitarians from libertarians) that can only be debated outside the theory. Ultimately, the theory of just market exchange is a module that can be incorporated into broader moral and political philosophies.

There is a second way in which the theory restricts subjectivity: It imposes strict consistency requirements to the moral judgments issued by a given observer. According to the theory, a sequence of moral judgments is consistent if and only if these judgements are rationalizable by the same disparity threshold function. If rationalization is impossible, the sequence of moral judgments must be considered arbitrary.

Our theory of just market exchange provides guidance for individual moral action. Nowhere does the theory say that its rules of conduct must be enforced by the state. Nevertheless, it is tempting to use the theory to justify state intervention in markets. Indeed, one can devise coherent political philosophies that incorporate the

---

[17] If a negotiation is fair, its degree of unfairness is zero. If a negotiation is unfair, its degree of unfairness is the difference between the weaker party's outside option and the minimum fair outside option. The minimum fair outside option is the increment that can be given to the weaker party such that the stronger party loses his unconscionable bargaining power.

theory of just market exchange plus other modules that demand the enforcement of its rules of conduct in specific contexts.

For example, the theory of just market exchange could be used to justify laws against price gouging, since after natural disasters the conditions for voluntariness fail: The breakdown of the supply chain creates temporary monopolies for the local suppliers of staples, while the desperate survivors are willing to pay for the staples inordinately high prices.

We advise against such policies.

Calculating the general, objective just price of a good is difficult; it may be impossible, without making heroic assumptions about value. We have proposed a partial solution, by giving a framework for arriving at the price a moral person would be allowed to charge (or offer; the case is symmetric). We have preserved the subjective element of value in the parameter of the disparity tolerance threshold, which is idiosyncratic. But given this basic idiosyncratic moral intuition we have advanced a system which allows an impartial observer to arrive at the implied moral maximum price.

References

Anomaly, J., G. Brennan, M. Munger, and G. Sayre-McCord. 2016. *Philosophy, politics, and economics: An anthology*. New York: Oxford University Press.

Brennan, J., and Peter J. 2015. *Markets Without Limits: Moral Virtues and Commercial Interests*. New York: Routledge Press.

Frankfurt, H. 1973. Coercion and moral responsibility. In *Essays on freedom of action*, ed. T. Honderich and K. Paul. London: Routledge.

Fried, C. 1981. *Contract as promise: A theory of contractual obligation*. Cambridge: Harvard University Press.

Gaus, G. 1990. *Value and Justification The Foundations of Liberal Theory*. New York: Cambridge University Press.

Guzman, R.A., and M. Munger. 2014. Euvoluntariness and just market exchange: Moral dilemmas from Locke's Venditio. *Public Choice* 158 (1–2): 39–49.

Hume, D. 1888. *A treatise of human nature*. Oxford: Clarendon Press.

Kumar, R. 2015. Risking and wronging. *Philosophy & Public Affairs* 43 (1): 27–51.

Levy, Yair. 2019. Constructivism and the problem of normative indeterminacy. *Journal of Value Inquiry* **(in press)**.

Locke, J. 2016. Venditio. In *Philosophy, Politics, and Economics: An Anthology*, ed. J. Anomaly, G. Brennan, M. Munger, and G. Sayre-McCord. New York: Oxford University Press

Lyons, D. 1975.Welcome threats and coercive offers. *Philosophy* 50 (194): 425–436.

McGregor, J. 1988. Bargaining advantages and coercion in the market. *Philosophy Research Archives* 14: 23–50.

Mizzoni, J. 2010. Recent work on evolution and social contract ethics. *The Journal of Value Inquiry* 44 (3): 377–388.

Munger, M.C. 2011. Euvoluntary or not, exchange is just. *Social Philosophy and Policy* 28 (2): 192–211.

Nagel, T. 1991. *Equality and partiality*. Oxford: Oxford University Press.

Olsaretti, S. 1998. Freedom, force and choice: Against the rights-based definition of voluntariness. *Journal of Political Philosophy* 6 (1): 53–78.

Olsaretti, S. 2004. *Liberty, desert and the market*. New York: Cambridge University Press.

O'neill, O. (1985). Between consenting adults. Philosophy of Public Affairs, 252-277.

Sample, R. 2003. *Exploitation: what It is and why it's wrong?*. Lanham: Rowman and Littlefield.

Sandel, M.J. 1998. What money can't buy: The moral limits of markets. *Tanner Lectures on Human Values* 21: 89–122.

Schmidtz, D. 2000. Diminishing marginal utility and egalitarian redistribution. *The Journal of Value Inquiry* 34 (2): 263–272.

Smith A (1776/1796) An Inquiry into the ature and auses of the ealth of ations. University of Chicago, Chicago.

Snyder, J. 2013. Exploitation and demeaning choices. *Politics, Philosophy & Economics* 12 (4): 345–360.

Snyder, J.C. 2008. Needs exploitation. *Ethical Theory and Moral Practice* 11 (4): 389–405.

Timmermann, J. 2005. Good but not required?—Assessing the demands of Kantian ethics. *Journal of Moral Philosophy* 2: 9–27.

Vrousalis, N. 2013. Exploitation, vulnerability, and social domination. *Philosophy & Public Affairs* 41 (2): 131–157.

Wells, M. 2017. Markets with some limits. *The Journal of Value Inquiry* 51 (4): 611–618.

Wertheimer, A. 1999. *Exploitation*. Princeton: Princeton University Press.

Weymark, J.A. 2014. Must one be an ogre to rationally prefer aiding the nearby to the distant needy? *Philosophy & Public Affairs* 42 (3): 230–252.

Zwolinski, M. 2008. The ethics of price gouging. *Business Ethics Quarterly* 18 (3): 347–378.

Zwolinski, M. 2009. Price gouging, non-worseness, and distributive justice. *Business Ethics Quarterly*, 19(2): 295–306.