



**BGPE Discussion Paper**

**No. 64**

**Fairness, Price Stickiness, and History  
Dependence in Decentralized Trade**

**Christian Korth  
Stefan Napel**

July 2008

---

ISSN 1863-5733

Editor: Prof. Regina T. Riphahn, Ph.D.  
Friedrich-Alexander-University Erlangen-Nuremberg  
© Christian Korth, Stefan Napel

---

# FAIRNESS, PRICE STICKINESS, AND HISTORY DEPENDENCE IN DECENTRALIZED TRADE

Christian Korth<sup>a</sup>

and

Stefan Napel<sup>a,★</sup>

July 31, 2008

## ABSTRACT

The paper investigates price formation in a decentralized market with random matching. Agents are assumed to have subdued social preferences: buyers, for example, prefer a lower price to a higher one but experience reduced utility increases below a reference price which serves as a common fairness benchmark. The strategic equilibrium reflects market fundamentals, but it is markedly less sensitive to the buyer-seller ratio near the fair price benchmark. Prices may be sticky around very different reference levels in markets with otherwise identical fundamentals. The implied history dependence turns out to be mitigated rather than exacerbated by friction.

**JEL codes:** C78; D49; D63; C91

**Keywords:** random matching; price stickiness; social preferences; history dependence; reference dependence;

<sup>a</sup> Dept. of Economics, University of Bayreuth, Universitätsstr. 40, 95447 Bayreuth, Germany;  
+49-921-55288-0/9 (phone/fax); christian.korth@uni-bayreuth.de; stefan.napel@uni-bayreuth.de

★ Corresponding author

## 1. INTRODUCTION

Fair behavior in bilateral bargaining situations has been one of the most extensively researched areas in microeconomics – experimentally and theoretically. Whether the robust evidence that people are concerned with fairness in bilateral bargaining (see, e.g., Camerer 2003) has implications for interaction on markets, however, remains unclear. Early studies by Kahneman et al. (1986), followed up by Kachelmeier et al. (1991) and Franciosi et al. (1995), have demonstrated that there are fairness effects on markets. However, they have been observed mainly as a transient phenomenon: the prices in the reported experiments typically approach the standard competitive equilibrium as time passes and unsustainable expectations are unwound. And many experimental studies have failed to find significant fairness effects on markets altogether. In their seminal paper, Fehr and Schmidt (1999, p. 818) summarize that

“... in competitive experimental markets with complete contracts, in which a well-defined homogenous good is traded, *almost all* subjects behave as if they are only interested in their material payoff. Even if the competitive equilibrium implies an extremely uneven distribution of the gains from trade, equilibrium is reached within a few periods.”

The experimental investigations referred to by Fehr and Schmidt mostly concern double auction settings, and are therefore difficult to interpret: the zero-intelligence trader results of Gode and Sunder (1993) have highlighted the strong built-in tendency of double auctions to reproduce the competitive equilibrium. Whether agents lose their concern for fairness when they interact in market environments or if their social preferences are simply overwhelmed by institutional structure is hard to discern. In fact, Brown et al. (2004) have recently shown that social preferences can play an important role in the presence of contractual incompleteness. Their experimental results complement empirical findings by Young and Burke (2001) which testify to the importance of custom and focal division rules.

Many market forms other than double auctions are of practical interest. The *search* or *matching markets* studied by Diamond and Maskin (1979), Mortensen (1982), or Rubinstein and Wolinsky (1985) are cases in point. They feature prominently in labor, real estate and monetary economics.<sup>1</sup> A key property of these markets is that they temporarily involve one-to-one interaction. This establishes a direct social relation and moreover creates leeway for prices to reflect the attitudes towards risk, delay, and possibly fairness of any two trade partners.

---

<sup>1</sup>See, e.g., Mortensen and Pissarides (1999) or Rogerson et al. (2005) for labor search models, Krainer and LeRoy (2002) or Albrecht et al. (2007) for housing markets, and Shi (2001) or Rocheteau and Wright (2005) for monetary applications.

This paper investigates the implications that social preferences have for the stationary strategic equilibrium of such a decentralized market. Agents are supposed to be averse to unfairly unfavorable as well as unfairly favorable deals in the spirit of Fehr and Schmidt (1999),<sup>2</sup> but we stay rather close to standard individualistic preferences: the negative weights on advantageous and disadvantageous deviations from what is considered as the fair benchmark are such that utility remains strictly increasing in own surplus share. And in contrast to the original Fehr-Schmidt model, the fair split need not automatically be a 50–50 division; any price between sellers' cost and buyers' willingness to pay may be the one which – for whatever reasons – is agents' reference point in a given market. This makes it possible to consider a more flexible notion of fairness than is usually done. It is in line with cognitive dissonance theory from psychology and the noteworthy experiments of Binmore et al. (1991) and Binmore et al. (1993), where subjects who were triggered to play different bargaining equilibria ended up considering very different surplus distributions as 'fair'.<sup>3</sup>

Apart from agents' concern with fairness, the investigated model is a simple version of the bargaining markets investigated by Rubinstein and Wolinsky (1985), Gale (1986a; 1986b) or Binmore and Herrero (1988): buyers and sellers are randomly divided into pairs. One randomly selected partner in each match suggests a deal. Rejection dissolves the match and agents wait to be rematched; successful traders leave the market and new ones enter. For simplicity, buyers and sellers are considered to be perfectly homogeneous. And all relevant information – most notably about agents' preferences – is assumed to be common knowledge.

Examples of real world markets in which agents can be thought of as interacting loosely as in our model include housing markets with approximately stationary landlord and tenant populations, similarly stationary labor markets with individual contracting, consumer-to-consumer sales of used cars, or direct procurement of differentiated commodities such as specialty tea, coffee, or wine. Agents in these markets may have an idea about the 'right', 'fair', or 'appropriate' rent, wage, or price which affects their subjective evaluation of a proposal (irrespective of informational imperfections which we completely leave out of our model). We show how already a rather weak concern for fairness can give rise to price stickiness. This fairness effect is predicted to persist over time for the considered market setting. It provides an explanation for price or wage rigidities that complements the traditional ones based on imperfect information, menu costs, long-term contracts, or money illusion (see, e.g., Woodford 2002 for the first, Fehr and Tyran 2001 for the latter, Golosov and Lucas 2007 for menu cost, and Ball and Mankiw 1994 or Taylor 1999 for general overviews).

---

<sup>2</sup>Prominent related models of social preferences include the ones by Bolton and Ockenfels (2000), Charness and Rabin (2002), Dufwenberg and Kirchsteiger (2004) or Falk and Fischbacher (2006).

<sup>3</sup>An endogenous notion of fairness also features prominently in Binmore's (1994, 1998) theory of distributional justice. It interprets the human sense of fairness as an efficient means to reach agreements quickly; it has evolutionary advantages only if it can adapt to changes in actual bargaining strength.

If the fairness benchmark itself is affected by agents' market experience, aversion to deviations from the given reference price gives rise to a continuum of possible self-confirming price conventions. This can explain distinct price levels in fundamentally identical markets. Interestingly, the scope of history dependence is larger, the less friction is associated with the rejection of an offer: friction turns out to erode the effects of fairness concern. Similarly, price stickiness is reduced rather than increased by friction.

We will now introduce the model. The ensuing strategic market equilibrium is studied in Section 3. We first discuss the market price's response to changes in the ratio of buyers to sellers for a given reference price, and then investigate the implications of endogenizing the latter. Section 4 concludes.

## 2. THE MODEL

Our basic market set-up mimics that of Rubinstein (1989, Model A): agents can trade a single indivisible good at discrete points of time indexed by  $t = 0, 1, 2, \dots$ . An agent is either a seller endowed with one unit of the good that she wishes to sell, or a buyer with at least one unit of disposable income. The sets of active buyers and sellers have cardinalities  $B$  and  $S$ , respectively.

In each period, agents are drawn randomly from the current population and matched with an agent of opposite type if there is one. For simplicity, the matching technology is assumed to involve no friction other than delay after the rejection of an offer, i.e., all agents on the short side of the market are matched with certainty.<sup>4</sup> Those on the long side of the market each have the same probability of being matched with any specific trade partner, independently of the history of play.

All sellers have a reservation price of zero and all buyers have a reservation price of one. After being matched, each couple engages in a bargaining session in which either the seller or the buyer is selected with equal probability to propose a price  $p \in [0, 1]$  that shall be paid by the buyer in exchange for the good. The quoted price amounts to a proposal of how to split the total monetary surplus of size one which is generated by a trade. The responding agent can accept or reject the offer. If it is rejected, then the two agents stop bargaining and both return to the set of active buyers or sellers. If the offer is accepted, then the trade is carried out and both agents leave the market. In this case, a new pair of seller and buyer is assumed to arrive to the market, i.e., the number of active agents in the market is presumed to be constant over time. Intuitively, in case of, e.g., a housing market this amounts to positing that new apartments are rented at a rate similar to the one at which agents vacate apartments and start searching for a new one.

---

<sup>4</sup>Assuming that agents on the short side of the market are matched with probability  $\kappa < 1$  would make the expressions below even more cumbersome, without qualitatively different findings.

All agents of type  $B$  or  $S$  (buyer or seller) are assumed to have identical von Neumann-Morgenstern utility functions which can reflect a fairness preference as proposed by Fehr and Schmidt (1999). In case that an agent never trades, her utility is zero. If an agent of type  $i \in \{S, B\}$  trades, then her utility  $u_i$  depends on the period  $t \geq 0$  in which the agreement is reached, the accepted price  $p \in [0, 1]$ , and also on an exogenous reference price  $p_{\text{ref}} \in [0, 1]$ . This reference price is assumed to be the same for every agent and meant to capture a common notion of the ‘fair price’ for the good, e.g., the average rent per square meter published in municipal housing statistics.<sup>5</sup>

Agents are assumed to be averse towards deviations from the reference price, i.e., a deviation of price and reference price diminishes utility. Positive and negative differences may be judged asymmetrically: deviations to one’s own favor (higher price for the seller, lower price for the buyer) are weighted by a parameter  $\alpha_i \geq 0$ , those to one’s disfavor by  $\beta_i \geq 0$ . This would in principle allow utility functions to have a global maximum at the reference price, leading to an equilibrium in which the agents always agree exactly on that price. To avoid this trivial case we impose that  $\alpha_i \in [0, 1)$ : a seller always prefers a higher price to a lower price, and a buyer always prefers a lower price to a higher one. Note that utility remains strictly monotonic in surplus as in Rubinstein and Wolinsky’s original setting, but agents will now prefer a sure payoff to a lottery between two prices with the same expected value if one price deviates from the reference price to their advantage and the other to their disadvantage.

Agents have stationary time preferences and discount future utility gains – for simplicity, by a common factor  $\delta \in (0, 1)$ . Overall, preferences are assumed to be given by

$$(1) \quad u_S(p, p_{\text{ref}}, t) = \delta^t (p - \alpha_S(p - p_{\text{ref}})^+ - \beta_S(p_{\text{ref}} - p)^+)$$

for sellers and

$$(2) \quad u_B(p, p_{\text{ref}}, t) = \delta^t (1 - p - \alpha_B(p_{\text{ref}} - p)^+ - \beta_B(p - p_{\text{ref}})^+)$$

for buyers where  $(x)^+ \equiv \max\{0, x\}$ . The utility functions are illustrated in Figure 1 (solid lines) for a trade in the current period. Note that the indicated kinks do not imply kinked aggregate demand or supply curves: demand is zero for prices which imply negative buyer utility and then jumps to  $B$ . Similarly, aggregate supply drops from  $S$  to zero.

The considered preferences coincide with those proposed by Fehr and Schmidt (1999) if  $p_{\text{ref}} = 0.5$  and fairness concern applies only to the two agents within a match, not those outside. We deem the restriction to one’s direct trade partner (as in Brown et al. 2004) a good starting point – in particular for markets with many buyers and sellers. As an extension of our analysis one might also consider agents who compare themselves with, e.g., the entire

---

<sup>5</sup>We will briefly point out the effects of allowing distinct reference prices amongst buyers and sellers in Section 4.

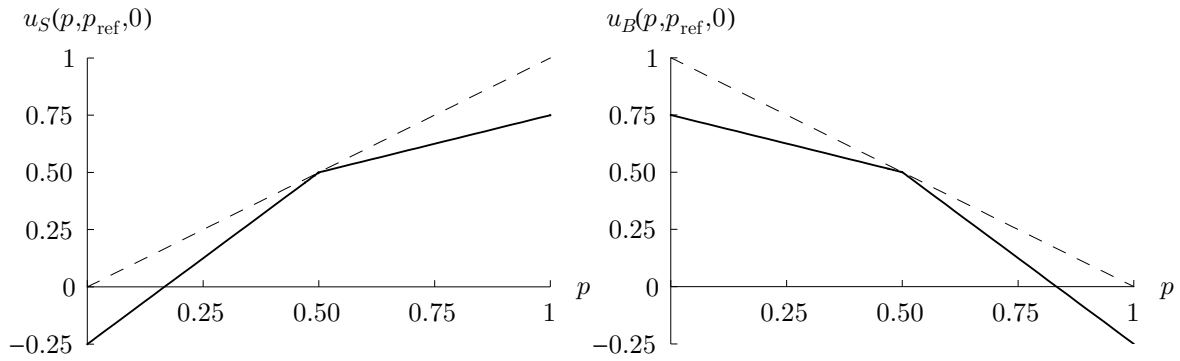


FIGURE 1. Kinked utility functions of seller and buyer with  $\alpha_i = \beta_i = 0.5$  and  $p_{\text{ref}} = 0.5$

population. As in Ewerhart's (2004) investigation of two-player Rubinstein bargaining, we allow other prices than 0.5 to be considered as fair. We will not construct any detailed model of the 'fair price' and initially consider it simply as given. We refer to Bolton et al. (2003) for a discussion of the determining roles played by past prices, competitor prices, and production costs, respectively.<sup>6</sup> However, we will later interpret  $p_{\text{ref}}$  as an endogenous fairness norm, as suggested by Binmore et al. (1993) and Binmore (1998), and require that its distance to the actual market outcome cannot exceed an exogenous threshold, which is interpreted as the maximal tolerated level of cognitive dissonance, in the long run.

### 3. ANALYSIS OF THE MODEL

We will consider only symmetric equilibria in which all buyers and sellers respectively use the same sequentially rational strategy. These strategies may condition agents' actions on the current period  $t$  but not the particular history of play or agents' identities. Denote the value (or expected indirect utility) of being a seller in period  $t$  by  $V_S(t)$ , and the value of being a buyer in period  $t$  by  $V_B(t)$ . Let  $p_B(t)$  be the price offered in equilibrium by the buyer if he is selected to make the proposal, and  $p_S(t)$  the price offered in equilibrium by the seller if she is to make the proposal. Both are selected with equal probabilities, and so  $p(t) \equiv \frac{1}{2}(p_S(t) + p_B(t))$  is the average price in the market in period  $t$ . When there is no danger of confusion we will suppress  $p_{\text{ref}}$  or  $t$  in agents' utility functions.

Whichever equilibrium arises,  $\delta V_S(t) + \delta V_B(t) < 1$  must hold because

$$u_S(p, p_{\text{ref}}, 0) + u_B(p, p_{\text{ref}}, 0) \leq 1$$

<sup>6</sup>Also see Xia et al. (2004) for a survey of psychological and sociological research on price fairness, and Köszegi and Rabin (2006) for a model of preferences which depend on an endogenous reference point determined by rational expectations about outcomes.

for all  $p, p_{\text{ref}}, \alpha, \beta \in [0, 1]$  and  $\delta < 1$ . Therefore, every match must end with a transaction: if this were not the case, then there would be a period  $t$  in which at least one proposer quotes a price that is rejected in equilibrium. Now the proposer could gain by making a deviating offer  $p$  that satisfies  $u_S(p, t) > \delta V_S(t+1)$  and  $u_B(p, t) > \delta V_B(t+1)$ : any sequentially rational responder strategy would entail acceptance of this offer.

In equilibrium the buyer will always offer exactly  $p_B = u_S^{-1}(\delta V_S(t+1))$  and the seller will always offer  $p_S = u_B^{-1}(\delta V_B(t+1))$  in order to render the respective responder indifferent; these offers will be accepted immediately. Assuming that sellers are on the short side of the market or both types are equally numerous ( $S \leq B$ ), the following equations must hold for all  $t$ :

$$(3) \quad u_S(p_B, t) = \delta V_S(t+1)$$

$$(4) \quad u_B(p_S, t) = \delta V_B(t+1)$$

$$(5) \quad \begin{aligned} V_S(t) &= \frac{1}{2}(u_S(p_B, t) + u_S(p_S, t)) \\ &= \frac{1}{2}(\delta V_S(t+1) + u_S(u_B^{-1}(\delta V_B(t+1)))) \end{aligned}$$

$$(6) \quad \begin{aligned} V_B(t) &= \frac{S}{B} \left[ \frac{1}{2}(u_B(p_S, t) + u_B(p_B, t)) \right] + \left(1 - \frac{S}{B}\right) [\delta V_B(t+1)] \\ &= \frac{S}{2B} [\delta V_B(t+1) + u_B(u_S^{-1}(\delta V_S(t+1)))] + \left(1 - \frac{S}{B}\right) [\delta V_B(t+1)] \end{aligned}$$

$$(7) \quad p(t) = \frac{1}{2} [p_B + p_S] = \frac{1}{2} [u_B^{-1}(\delta V_B(t+1)) + u_S^{-1}(\delta V_S(t+1))]$$

Analogous expressions characterize the strategic market equilibrium if buyers are on the short side of the market ( $S > B$ ). We refer to Rubinstein (1989) for details.

**3.1. Equilibrium Price.** Denote the fraction of buyers in the market by  $b \equiv \frac{B}{S+B} \in (0, 1)$ . Equations (3)–(7) have a unique and time-independent solution  $p(t) \equiv p^*$  with three cases that depend on the level of  $p_{\text{ref}}$  and the other parameters of the model:

$$(8) \quad p^* = \begin{cases} \frac{[b+(2b(1-\delta)+\delta)\beta_S]p_{\text{ref}}(1-\alpha_B)+b(1-p_{\text{ref}})(1+\beta_S)}{[2b(1-\delta)+\delta](1-\alpha_B)(1+\beta_S)} & \text{if } p_{\text{ref}} \geq \bar{p} \quad (a) \\ \left[ \begin{array}{l} \delta p_{\text{ref}} [\beta_B + (2-\delta)(1+\beta_B)\beta_S + \delta\alpha_S + \alpha_B(1-\alpha_S\delta)] \\ + b[2+2\beta_B p_{\text{ref}} + \beta_S(2-\delta)(1-\delta+(1-2\delta+\beta_B(2-3\delta))p_{\text{ref}}) \\ - \delta(2+\alpha_B p_{\text{ref}} + 3\beta_B p_{\text{ref}} + \alpha_S(1-\delta-p_{\text{ref}} + (2-\alpha_B)\delta p_{\text{ref}}))] \end{array} \right] & \text{if } \underline{p} < p_{\text{ref}} < \bar{p} \quad (b) \\ / \left[ \begin{array}{l} 4b(1+\beta_B)(1+\beta_S) + 2(1-4b)(1+\beta_B)(1+\beta_S)\delta \\ - \delta^2[2-\alpha_S(1-b) - 4b + \alpha_B(\alpha_S-1)(1-b) \\ + (1-3b)(\beta_B + \beta_S + \beta_B\beta_S)] \end{array} \right] & \\ \frac{[b-(2b(1-\delta)+\delta)\alpha_S]p_{\text{ref}}(1+\beta_B)+b(1-p_{\text{ref}})(1-\alpha_S)}{[2b(1-\delta)+\delta](1-\alpha_S)(1+\beta_B)} & \text{if } p_{\text{ref}} \leq \underline{p} \quad (c) \end{cases}$$

with

$$\bar{p} = \frac{b(2-\delta)(1+\beta_S)}{b(2-\delta)(1+\beta_S) + (1-b)\delta(1-\alpha_B)}$$



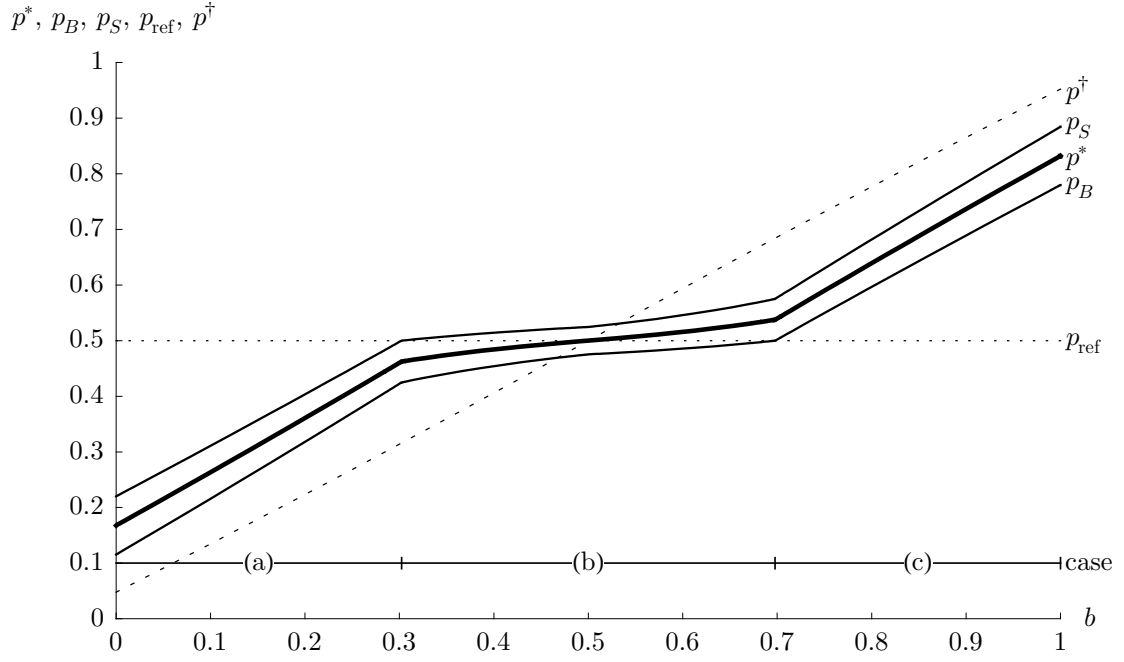


FIGURE 2. Equilibrium prices for  $\alpha_i = \beta_i = 0.3$ ,  $p_{\text{ref}} = 0.5$  and  $\delta = 0.95$

and

$$\underline{p} = \frac{b\delta(1-\alpha_S)}{(2b-3b\delta+\delta)(1+\beta_B)+b\delta(1-\alpha_S)}.$$

This simplifies somewhat if all fairness parameters are equal, i.e., for  $\alpha_S = \alpha_B = \beta_S = \beta_B \equiv \gamma$  we have

$$p^* = \begin{cases} \frac{b(\gamma+1)-\gamma(2b\gamma+(2b-1)(1-\gamma)p)p_{\text{ref}}}{(1-\gamma^2)(2b(1-p)+p)} & \text{if } p_{\text{ref}} \geq \bar{p} \quad (\text{a}) \\ \frac{\gamma p(-p\gamma+\gamma+2)p_{\text{ref}}+b(-p+\gamma((1-p)^2+2(\gamma(1-p)^2-2p+1)p_{\text{ref}})+1)}{p((\gamma+1)^2-(\gamma^2+1)p)+2b((\gamma+1)^2(1-p)^2-\gamma p^2)} & \text{if } \underline{p} < p_{\text{ref}} < \bar{p} \quad (\text{b}) \\ \frac{b(2(p-\gamma(1-p))p_{\text{ref}}\gamma-\gamma+1)-\gamma(\gamma+1)pp_{\text{ref}}}{(1-\gamma^2)(2b(1-p)+p)} & \text{if } p_{\text{ref}} \leq \underline{p} \quad (\text{c}) \end{cases}$$

$$\text{with } \bar{p} = \frac{b(2-\delta)(1+\gamma)}{2b(1-\delta+\gamma)+\delta(1-\gamma)} \quad \text{and} \quad \underline{p} = \frac{b\delta(1-\gamma)}{2b(1-\delta-2\delta\gamma+\gamma)+\delta(1+\gamma)}.$$

Figure 2 visualizes this term.<sup>7</sup> The thin solid lines indicate the proposer-specific prices  $p_S$  and  $p_B$  which give rise to the average market price  $p^*$ . Their respective intersections with the dotted reference price line coincide with the case boundaries.

<sup>7</sup>The price formulas shown above apply when  $S \leq B$ , i.e.,  $b = \frac{B}{B+S} \geq 0.5$ . The expressions in case of  $b < 0.5$  are symmetric (see the left half of the Figure 2). The above term for case (a) becomes relevant under  $S \leq B$  only if  $p_{\text{ref}}$  is very high. In particular, Figure 2 actually shows the price arising in case (a) of the solution for  $S > B$ .

In the first case, (a), the fair price  $p_{\text{ref}}$  is so high that  $p_B \leq p_S \leq p_{\text{ref}}$ . This means that both the seller and the buyer would offer a price below the fair price in equilibrium. In the last case, (c), the fair price  $p_{\text{ref}}$  is so low that  $p_S \geq p_B \geq p_{\text{ref}}$ , and both seller and buyer would offer a price above the fair price in equilibrium. The intermediate case, (b), with  $p_B < p_{\text{ref}} < p_S$  entails that if the seller is to propose a price, she suggests one above the fair price, while if the buyer is to propose a price, this price will be below the fair price.<sup>8</sup>

Figure 2 illustrates the key effect of fairness concern in a matching market: in a neighborhood of  $p_{\text{ref}}$ , the equilibrium price is much less sensitive to changes in the buyer-seller ratio than what would be the case without social preferences, i.e., the price is stickier. More formally, let

$$p^\dagger \equiv p^* \Big|_{\alpha_i = \beta_i = 0} = \frac{b}{2b(1-\delta) + \delta}$$

be the associated equilibrium price for standard preferences without fairness concern (broken line in Figure 2).<sup>9</sup> We will focus on markets – as characterized by the preference parameters  $\alpha_i, \beta_i, \delta$  and  $p_{\text{ref}}$  – for which  $p_{\text{ref}}$  arises as the equilibrium price for *some* fraction of buyers in the market. We denote this buyer share by  $\tilde{b}$ , i.e.,  $p^* \Big|_{b=\tilde{b}} = p_{\text{ref}}$ .<sup>10</sup> Then the following is true:

**PROPOSITION 1.** *The market price is less sensitive to changes in the fraction  $b$  of buyers in the market in a neighborhood of  $\tilde{b}$  if  $\max\{\alpha_B, \alpha_S, \beta_B, \beta_S\} > 0$  than if agents have no fairness concern:*

$$0 < \frac{\partial p^*}{\partial b} \Big|_{b=\tilde{b}} < \frac{\partial p^\dagger}{\partial b} \Big|_{b=\tilde{b}}$$

*Proof.* Buyer shares  $b$  in a sufficiently small neighborhood of  $\tilde{b}$  pertain to case (b): sellers propose a price above  $p_{\text{ref}}$  and buyers propose a price below  $p_{\text{ref}}$ . The derivatives of  $p^*$  and  $p^\dagger$  with respect to  $b$  and evaluated at  $\tilde{b}$  are therefore

$$\frac{\partial p^*}{\partial b} \Big|_{b=\tilde{b}} = \left\{ \begin{array}{l} \left[ (\delta-1)(2p_{\text{ref}}(\beta_B + \beta_S + 2) + \beta_S(\delta-2) + \alpha_S\delta - p_{\text{ref}}(-\alpha_B + \alpha_S + 3\beta_B + \beta_S + 4)\delta - 2)^2 \right] \\ / \left[ \begin{array}{l} \delta(\beta_S(\delta-2) + \alpha_S\delta - 2)(2(\beta_B + 1)(\beta_S + 1) - (\alpha_B(\alpha_S - 1) - \alpha_S + \beta_B + \beta_S\beta_S + \beta_S + 2)\delta) \\ + p_{\text{ref}}((\beta_B + 1)\beta_S(\delta-2) - (\alpha_S + \beta_B)\delta + \alpha_B(2\beta_B(\delta-1) + \alpha_S\delta + \delta - 2)) \end{array} \right] \end{array} \right\}$$

and

$$\frac{\partial p^\dagger}{\partial b} \Big|_{b=\tilde{b}} = \frac{((\delta-2)(p_{\text{ref}}-1)\beta_S + (3\delta-2)p_{\text{ref}}\beta_B + \delta(p_{\text{ref}}-1)\alpha_S + 4(\delta-1)p_{\text{ref}} + 2)^2}{\delta(2 - \delta\alpha_S + (\delta-2)(p_{\text{ref}}-1)\beta_S - 2p_{\text{ref}}\alpha_B + \delta p_{\text{ref}}(\alpha_B + \alpha_S + \beta_B))^2},$$

<sup>8</sup>No further cases arise because  $p_S \geq p_B$  and  $0 \leq p \leq \bar{p} \leq 1$ . The case boundaries can, of course, also be expressed in terms of  $b$  rather than  $p_{\text{ref}}, p$ , and  $\bar{p}$ : case (a) holds if  $b \leq \underline{b}$ , case (b) if  $\underline{b} < b < \bar{b}$ , and case (c) if  $\bar{b} \leq b$  with  $\underline{b} = \frac{\delta p_{\text{ref}}(1-\alpha_B)}{(2-\delta)(1-p_{\text{ref}})(1+\beta_S) + \delta p_{\text{ref}}(1-\alpha_B)}$  and  $\bar{b} = \frac{\delta p_{\text{ref}}(1+\beta_B)}{(3\delta-2)p_{\text{ref}}(1+\beta_B) + \delta(1-p_{\text{ref}})(1-\alpha_S)}$ .

<sup>9</sup>Again, we only provide explicit expressions for  $S \leq B$ , but analogous statements apply when  $S > B$ .

<sup>10</sup>A fraction  $\tilde{b}$  of buyers which allows  $p_{\text{ref}}$  to be an equilibrium price need not exist, e.g., when  $\delta$  is small and  $p_{\text{ref}}$  is either very high or low. However, one can check that for any given values of  $\alpha_i, \beta_i$  and  $\delta$ , there exist an interval of prices  $p$  which would respectively arise in equilibrium for some  $\tilde{b} \in (0, 1)$  provided that  $p_{\text{ref}} = p$ .

respectively. The range  $[p_{\text{ref}}, \bar{p}_{\text{ref}}]$  of reference prices  $p_{\text{ref}}$  for which  $\tilde{b}$  exists (see fn. 10) can be identified by solving

$$p^* \Big|_{b=\frac{1}{2}, p_{\text{ref}}=p_{\text{ref}}} = p_{\text{ref}} \quad \text{and} \quad p^* \Big|_{b=1-, p_{\text{ref}}=\bar{p}_{\text{ref}}} = \bar{p}_{\text{ref}}$$

because  $p^*$  is increasing in  $b$ . It can then be checked numerically that the claimed inequality holds for any  $p_{\text{ref}} \in [p_{\text{ref}}, \bar{p}_{\text{ref}}]$ .  $\square$

The driving force for the decreased sensitivity of prices near  $p_{\text{ref}}$ <sup>11</sup> is the local increase in concavity of the agents' utility function: price proposals by buyers and sellers that are, respectively, below and above the reference price ex ante amount to a lottery, towards which both agents are risk averse. In order to avoid this lottery, they are willing to accept a proposal (and thereby generate a safe return) which is closer to  $p_{\text{ref}}$  than it would be for agents whose utility is linear everywhere. Importantly, decreased price sensitivity would also obtain if the baseline utility were already strictly concave rather than linear in price below and above  $p_{\text{ref}}$ . Moreover, one could also assume an everywhere differentiable utility function without changing our qualitative findings. For example,

$$\tilde{u}_S(p, p_{\text{ref}}, t) = \delta^t \left( p + \frac{1}{2}(p - p_{\text{ref}}) \left( -\alpha_S + \beta_S - (\alpha_S + \beta_S) \tanh(c(p - p_{\text{ref}})) \right) \right)$$

is smooth and produces the same stickiness as  $u_S$  for large  $c > 0$  (it converges uniformly to  $u_S$  for  $c \rightarrow \infty$ ). The aspect of Fehr-Schmidt preferences which is crucial in our context is that fairness concern increases risk aversion in a neighborhood of the reference price.

When agents have *no* concern for fairness, the market price responds more to changes in the buyer-seller ratio, the closer  $\delta$  is to unity, i.e.,

$$\frac{\partial^2 p^+}{\partial b \partial \delta} = \frac{\delta(2b - 1) + 2b}{(2b(1 - \delta) + \delta)^3} > 0,$$

and  $\lim_{\delta \rightarrow 1} \frac{\partial p^+}{\partial b} = 1$ . So it is somewhat surprising that the stickiness or decreased sensitivity of prices near  $p_{\text{ref}}$  in case of  $\alpha_i > 0$  or  $\beta_i > 0$  is larger, the less friction there is: when agents *have* concern for fairness,  $p^*$  increases in  $b$  by less, the larger is  $\delta$ . This even gives rise to a flat price  $p^* \equiv p_{\text{ref}}$  for all buyer-seller ratios pertaining to case (b) in the limit  $\delta \rightarrow 1$ :

**PROPOSITION 2.** *Suppose that  $\max\{\alpha_B, \alpha_S, \beta_B, \beta_S\} > 0$ . Then, first, there exists some  $\delta' \in (0, 1)$  such that the market price's sensitivity to changes in the fraction  $b$  of buyers in a neighborhood of  $\tilde{b}$  is decreasing in  $\delta$  for any  $\delta > \delta'$ . Second, the price sensitivity  $\frac{\partial p^*}{\partial b} \Big|_{b=b'}$  converges to zero as  $\delta \rightarrow 1$  for any  $b'$  pertaining to case (b) of equation (8) (in particular, for  $b = \tilde{b}$ ).*

<sup>11</sup>Whilst the strict inequalities in Proposition 1 obviously hold on an entire interval around  $\tilde{b}$ , sensitivity of  $p^*$  with respect to the buyer-seller ratio need not be reduced at *every*  $b$ : it may locally be greater than for standard preferences near the boundary between cases (a) and (b), or (b) and (c). The average slope of  $p^*$  over the range of buyer-seller ratios pertaining to case (b) is, however, always smaller for positive  $\alpha_i$  or  $\beta_i$  than for  $\alpha_i = \beta_i = 0$ .

*Proof.* The cross-derivative  $\frac{\partial^2 p^*}{\partial b \partial \delta}$  is very unwieldy, but continuous in  $\delta$ . It therefore suffices to evaluate its limit as  $\delta \rightarrow 1$  in order to prove the first claim. One obtains<sup>12</sup>

$$\lim_{\delta \rightarrow 1} \left( \frac{\partial^2 p^*}{\partial b \partial \delta} \Big|_{b=\bar{b}} \right) = - \frac{(1 - p_{\text{ref}})(2 - \alpha_S + \beta_S)}{(b - 1)^2(\alpha_B + \alpha_S - \alpha_B \alpha_S + \beta_B + \beta_S + \beta_B \beta_S)} < 0.$$

To see that the second claim is true, note simply that  $\lim_{\delta \rightarrow 1} p^* = p_{\text{ref}}$  wherever case (b) applies.  $\square$

The economic intuition for this comparative static result is connected to the different response, as  $\delta$  increases, of the bargaining advantages conferred by (i) being on the short side of the market and (ii) currently being in the proposer position. To see this, first review the case *without* fairness concern. For concreteness, consider a situation in which buyers are on the short side of the market (i.e.,  $b < 0.5$ ). Parameter  $b$  directly affects sellers' continuation values, which are associated with rejecting an offer: if the current price is turned down, the seller remains unmatched in the next period with probability  $1 - b$ . This creates a bargaining advantage for buyers relative to sellers. Because it varies in  $b$ , so does  $p^*$ . The effect of this advantage is, however, attenuated by the equal distribution of another bargaining advantage: the right to make the first proposal. The value of this temporary 'monopoly position' is decreasing in  $\delta$ . So bargaining outcomes become increasingly driven only by the short sides' matching advantage. It is true that also this advantage becomes less important as  $\delta$  increases: waiting for a rematching bothers sellers less and less in absolute terms. But the asymmetric matching advantage turns out to vanish more slowly than the symmetrically distributed proposer advantage. Hence, prices become more responsive to the buyers' matching advantage, as parameterized by  $b$ .

*With* fairness concern, the bargaining advantage provided by the right to make the next proposal becomes more pronounced: sellers still hate accepting a low price, but now they are even less attracted by the lottery of getting a more decent price (above  $p_{\text{ref}}$ ) or facing the same proposal (below  $p_{\text{ref}}$ ) again next period than without fairness concern. An analogous statement applies to buyers. Critically, this (still symmetrically distributed) proposer advantage no longer vanishes as  $\delta \rightarrow 1$ : locally risk averse responders will prefer to accept a price near the benchmark even though there is the chance to almost immediately make a counteroffer. Agents on the long side must, of course, expect a somewhat greater delay until they get to make an offer after rejecting a proposal, and so their continuation value is still lower, the more asymmetric are both population sizes. However, as was also true without fairness concern, not being rematched immediately becomes less and less of a concern for large  $\delta$ . In the limit, as  $\delta \rightarrow 1$ , the matching advantage becomes negligible relative to the discounting-invariant part of the advantage enjoyed by any agent – from the short side or the long side of the market – who has been drawn as the current proposer.

<sup>12</sup>Again, we focus on the case of  $b \geq \frac{1}{2}$ , with an analogous expression applying to  $b < \frac{1}{2}$ .

**3.2. Self-confirming Price Conventions.** The reference price was so far treated as exogenously given by, e.g., the equal split considered by Fehr and Schmidt (1999), official rental statistics, or perhaps a recommended retail price. In line with a now sizeable economic literature, we presumed that agreements which are not in line with this benchmark entail a subjective penalty. That agents who consider themselves as fair persons attach an additional cost to concluding unfair deals also broadly fits psychologists' view. However, psychological research on *cognitive dissonance* in the tradition of Festinger (1957) (see Nail et al. 2004 or Cooper 2007 for recent overviews and, e.g., Akerlof and Dickens 1982 for an economic perspective on the literature) suggests that agents who stoically bear these costs are not the only possibility. An agent may rather adjust his notion of fairness, and this is particularly likely if deals are not even on average in line with her or his current notion of fairness.<sup>13</sup> We will therefore consider the possibility of the reference price being *endogenous* and at least partially responsive to economic fundamentals. The reference price could, for instance, plausibly arise as an average of past prices such as official rental statistics or depend in a more complex way on history. It is meant to capture a broadly shared notion of what constitutes a fair deal, which will plausibly be adapted if it should persistently and noticeably fall out of line with agents' actual experience.

This variation of the above baseline model fits nicely with Ken Binmore's (1994, 1998) theory of distributional justice and the still intriguing experimental findings of Binmore et al. (1993). This study investigates the role of focal points in bargaining and highlights that subjects' notion of a fair division is no universal constant but rather quite path-dependent. Specifically, Binmore et al. first made distinct surplus divisions focal to their subjects by having them bargain with different computer programs in the initial stage of the experiment, one program insisting on 50:50 splits, another one on 75:25 splits, and so on. The subjects switched between the proposer and responder roles in order to become familiarized with both perspectives. Then participants who had played the same program were matched and bargained with each other. The result was that subjects – in fundamentally identical bargaining situations – agreed on different surplus divisions depending on which divisions they had been conditioned to in the initial stage: the bargaining outcomes clustered around the respective induced focal division. And, more surprisingly, most subjects ended up describing the agreements that they had reached as particularly fair. What participants perceived as fair reflected their respective payoff experience (also see Binmore et al. 1991).

Adjusting one's notion of a fair deal in avoidance of prolonged cognitive dissonance is likely to be subjectively costly itself (see Konow 2000 for a theory of decision making which investigates this assumption in detail). It is worthwhile to incur the adjustment cost only

---

<sup>13</sup>This is analogous, for instance, to car drivers who are outraged by the common 5–10% gasoline price increase during public holidays with peak demand, but then come to 'accept' an altogether higher price level within a few months (think of recent increases above \$3.5/gallon or €1.4/litre).

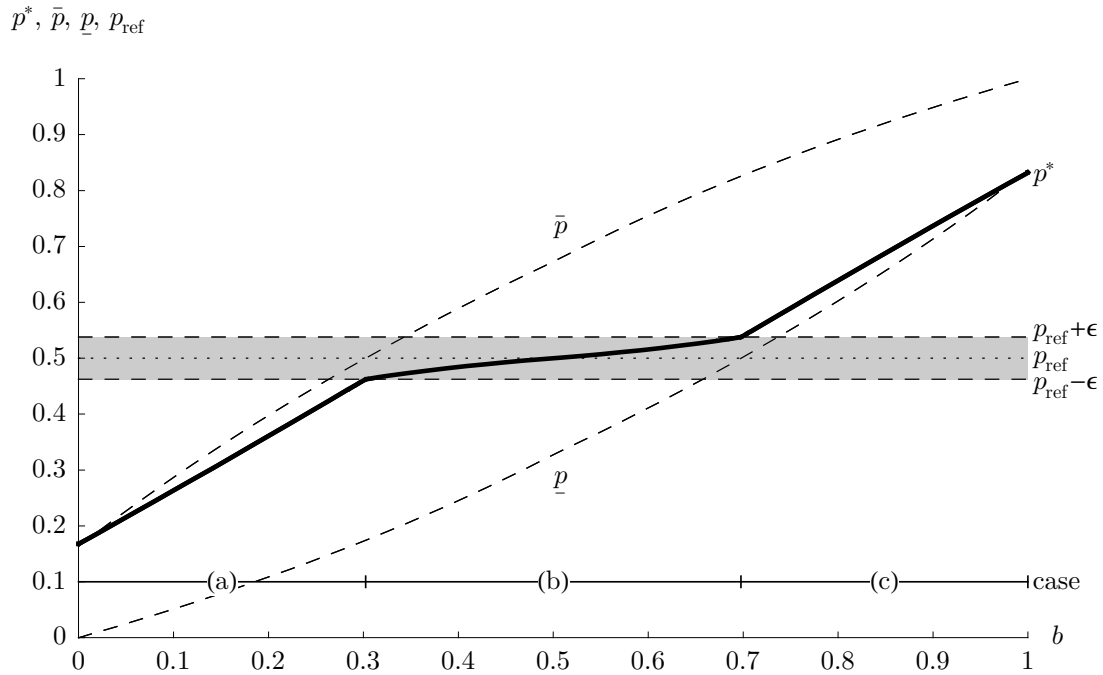


FIGURE 3. Self-confirming price convention  $p_{\text{ref}} = 0.5$  with  $\alpha_i = \beta_i = 0.3$  and  $\delta = 0.95$

if the dissonance cost associated with the conclusion of deals that are unfair according to the current fairness standard is even higher, i.e., if in our context the gap between  $p_{\text{ref}}$  and the average equilibrium price  $p^*$  is large enough. Denote by  $\epsilon \geq 0$  the size of the gap between the current reference price and the market price that agents would still tolerate rather than adapting their fairness benchmark in line with cognitive dissonance theory. In other words, agents are assumed to abandon the current reference price  $p_{\text{ref}}$  at some point in time if  $p^* \notin [p_{\text{ref}} - \epsilon, p_{\text{ref}} + \epsilon]$ . Such prices  $p_{\text{ref}}$  will be referred to as  $\epsilon$ -unstable. In contrast, a reference price  $p_{\text{ref}}$  which gives rise to an average equilibrium price  $p^*$  inside the respective  $\epsilon$ -neighborhood of  $p_{\text{ref}}$  would not be changed; it will be referred to as an  $\epsilon$ -stable price convention. We will assume the cost of adjustment which is indirectly measured by parameter  $\epsilon$  to be small but positive, i.e., we consider  $\epsilon > 0$  in the following.

An  $\epsilon$ -stable price  $p_{\text{ref}}$  is *self-confirming* in the following sense: agents regard  $p_{\text{ref}}$  as fair, this gives rise to an average market price in an  $\epsilon$ -neighborhood of  $p_{\text{ref}}$ , and this in turn confirms  $p_{\text{ref}}$  as agents' reference price. As illustrated in Figure 3, the property of a given reference price to be self-confirming can be robust to quite sizeable fluctuations in the buyer-seller ratio. It can be checked that if the tolerable gap between the fair and the equilibrium price is such that

$$(9) \quad \epsilon \geq \epsilon_{\min}(p_{\text{ref}}) \equiv \frac{1 - \delta}{\delta} \cdot \max \left\{ \frac{p_{\text{ref}}}{1 - \alpha_S}, \frac{1 - p_{\text{ref}}}{1 - \alpha_B} \right\},$$

then *any* buyer-seller ratio for which the seller and buyer prices are respectively above and below  $p_{\text{ref}}$ , i.e., any market for which case (b) applies, makes the given  $p_{\text{ref}}$  self-confirming.<sup>14</sup> Fairness concern may thus not only give rise to reduced short-run sensitivity of prices, but it can cause prices to be sticky, i.e., to be insensitive to variations of  $b$  in the long run, too.

One can also look at Figure 3 from an alternative perspective. Suppose that the buyer-seller ratio is fixed at some level within case (b) and one observes an equilibrium price  $p^* \in [0.5 - \epsilon; 0.5 + \epsilon]$  as depicted. Then  $p_{\text{ref}} = 0.5$  would be an  $\epsilon$ -stable price convention for the given buyer-seller ratio. However, it is not the only one. If, for instance,  $\epsilon$  is so large as to satisfy inequality (9) for  $p_{\text{ref}} = \underline{p}$  and for  $p_{\text{ref}} = \bar{p}$ , then *any* reference price  $\tilde{p}_{\text{ref}} \in [\underline{p}, \bar{p}]$  would be  $\epsilon$ -stable: the monotonicity of the two alternative determinants of  $\epsilon_{\min}(\cdot)$  in  $p_{\text{ref}}$  (namely,  $p_{\text{ref}}/(1 - \alpha_S)$  and  $(1 - p_{\text{ref}})/(1 - \alpha_B)$ ) implies that  $\epsilon \geq \epsilon_{\min}(\tilde{p}_{\text{ref}})$  holds for any  $\underline{p} \leq \tilde{p}_{\text{ref}} \leq \bar{p}$ . Therefore – recalling  $\epsilon_{\min}$ 's construction – any such  $\tilde{p}_{\text{ref}}$  would be self-confirming. Smaller thresholds  $\epsilon$  for the subconscious ‘correction’ of cognitive dissonance, of course, imply a smaller range of  $\epsilon$ -stable reference prices. But there must still exist an entire interval of self-confirming price conventions:

**PROPOSITION 3.** *For any  $b \in (0, 1)$  and  $\epsilon > 0$  there exists a non-empty interval  $[\underline{p}, \check{p}] \subseteq [\underline{p}, \bar{p}]$  such that any  $p \in [\underline{p}, \check{p}]$  would constitute a self-confirming reference price, i.e.,  $p^*|_{p_{\text{ref}}=p} \in [p - \epsilon, p + \epsilon]$ .*

*Proof.* Note that for any  $b \in (0, 1)$ , there exists a unique  $p^0 \in [\underline{p}, \bar{p}]$  which is 0-stable, i.e.,  $p^*|_{p_{\text{ref}}=p^0} = p^0$ . This can be evaluated as

$$p^0 = \frac{b(\beta_S(\delta - 2) + \alpha_S\delta - 2)}{(\alpha_B - \beta_B - 2 + b(4 - \alpha_B + \alpha_S + 3\beta_B + \beta_S))\delta - 2b(2 + \beta_B + \beta_S)}$$

from equation (8) case (b). It then follows directly from the continuity of  $p^*$  in  $p_{\text{ref}}$  that  $p^*|_{p_{\text{ref}}=p} \in [p - \epsilon, p + \epsilon]$  is satisfied for any  $p$  in a sufficiently small neighborhood of  $p^0$ .  $\square$

The only situation which does not create any long-run multiplicity of fair price conventions would be the one in which adjustments of agents’ fairness notion entails no costs at all and which therefore involves  $\epsilon = 0$ . Leaving this possibility of a perfectly flexible (and then perhaps vacuous) sense of fairness aside, there always exists a *continuum of  $\epsilon$ -stable price conventions* for the same market environment. This is an effect of fairness concern that comes on top of reduced short and long-run sensitivity of prices regarding changes in the buyer-seller ratio, as discussed earlier. Thinking of independent housing markets in city A and city B as an example, the observed price levels could differ by as much as  $\bar{p} - \underline{p}$  even though both have exactly the same fundamentals (identical marginal costs and marginal products of housing, same numbers of active buyers and sellers, identical parameters  $\alpha_i$ ,  $\beta_i$  and  $\delta$ ). The

<sup>14</sup>Given that  $p^*$  is monotone in  $b$ , the potential gap between  $p^*$  and  $p_{\text{ref}}$  for case (b) is largest at the case boundaries. So the expression for  $\epsilon_{\min}(p_{\text{ref}})$  follows from evaluating  $|p_{\text{ref}} - p^*|$  at  $\underline{p}$  and  $\bar{p}$ .

only difference would be that, for instance, because the collection of rental statistics initially had a bias in one but not the other community, cities A and B have different reference prices. These need not necessarily have been self-confirming from the beginning: a very high initial reference rent in city A will presumably give rise to a relatively high initial market price and a long-run equilibrium  $p_A^*$  near  $\bar{p}$ ; a low initial one in city B would give rise to  $p_B^*$  near  $\underline{p}$ .<sup>15</sup>

Echoing the comparative static findings in Proposition 2, the possible history dependence caused by fairness concern does not vanish as  $\delta \rightarrow 1$ ; rather, the range of self-confirming reference prices becomes larger as  $\delta$  increases:

PROPOSITION 4. For any given  $\epsilon > 0$ , the size of the interval  $[\underline{p}, \check{p}] \subseteq [\underline{p}, \bar{p}]$  such that  $p^*|_{p_{\text{ref}}=p} \in [p - \epsilon, p + \epsilon]$  holds for every  $p \in [\underline{p}, \check{p}]$  is increasing in  $\delta$ , i.e.,

$$\frac{\partial(\check{p} - \underline{p})}{\partial\delta} > 0.$$

Furthermore,  $[\underline{p}, \check{p}]$  converges to  $[\underline{p}, \bar{p}]$  as  $\delta \rightarrow 1$ .

*Proof.* The second part of the proposition follows directly from the fact that

$$\lim_{\delta \rightarrow 1} p^* = p_{\text{ref}}$$

for any  $p_{\text{ref}} \in [\underline{p}, \bar{p}]$ .

The endpoints of interval  $[\underline{p}, \check{p}]$  can be computed by solving  $p^*|_{p_{\text{ref}}=\check{p}} \equiv \check{p} - \epsilon$  for  $\check{p}$  and  $p^*|_{p_{\text{ref}}=\underline{p}} \equiv \underline{p} + \epsilon$  for  $\underline{p}$  in case (b) of equation (8). One obtains

$$\check{p} = \left\{ \begin{array}{l} \left[ \begin{array}{l} \delta(2(1+\beta_B)(1+\beta_S) - (2+\alpha_B(-1+\alpha_S) - \alpha_S + \beta_B + \beta_S + \beta_B\beta_S)\delta)\epsilon + b(2+4(1+\beta_B)\epsilon - \delta(2+\alpha_S \\ + 8(1+\beta_B)\epsilon) + \delta^2(\alpha_S + (4+\alpha_B(-1+\alpha_S) - \alpha_S + 3\beta_B)\epsilon) + \beta_S(-2+\delta)(-1+\delta+(1+\beta_B)(-2+3\delta)\epsilon)) \end{array} \right] \\ \left[ \begin{array}{l} (-1+\delta)(-2b(2+\beta_B+\beta_S) + (-2+\alpha_B-\beta_B+b(4-\alpha_B+\alpha_S+3\beta_B+\beta_S))\delta) \end{array} \right] \end{array} \right\}$$

and

$$\underline{p} = \left\{ \begin{array}{l} \left[ \begin{array}{l} (\delta(-2(1+\beta_B)(1+\beta_S) + (2+\alpha_B(-1+\alpha_S) - \alpha_S + \beta_B + \beta_S + \beta_B\beta_S)\delta)\epsilon + b(2-4(1+\beta_B)\epsilon - \beta_S(-2+\delta) \\ (1-\delta+(1+\beta_B)(-2+3\delta)\epsilon) + \delta(-2+8(1+\beta_B)\epsilon + (-4+\alpha_B-3\beta_B)\delta\epsilon + \alpha_S(-1+\delta(1+\epsilon-\alpha_B\epsilon)))) \end{array} \right] \\ \left[ \begin{array}{l} (-1+\delta)(-2b(2+\beta_B+\beta_S) + (-2+\alpha_B-\beta_B+b(4-\alpha_B+\alpha_S+3\beta_B+\beta_S))\delta) \end{array} \right] \end{array} \right\}.$$

The size of the interval therefore equals

$$\check{p} - \underline{p} = \left\{ \begin{array}{l} \left[ \begin{array}{l} 2\epsilon(4b(1+\beta_B)(1+\beta_S) - 2(-1+4b)(1+\beta_B)(1+\beta_S)\delta + (-2+\alpha_B+\alpha_S \\ -\alpha_B\alpha_S - \beta_B - \beta_S - \beta_B\beta_S + b(4+\alpha_B(-1+\alpha_S) - \alpha_S + 3\beta_S + 3\beta_B(1+\beta_S)))\delta^2) \end{array} \right] \\ \left[ \begin{array}{l} (-1+\delta)(-2b(2+\beta_B+\beta_S) + (-2+\alpha_B-\beta_B+b(4-\alpha_B+\alpha_S+3\beta_B+\beta_S))\delta) \end{array} \right] \end{array} \right\}.$$

The partial derivative  $\frac{\partial(\check{p}-\underline{p})}{\partial\delta}$  is again quite unwieldy, but can be shown to be strictly positive for any  $b$  pertaining to case (b) given the restrictions on  $\alpha_i$ ,  $\beta_i$  and  $\delta$  introduced above.  $\square$

<sup>15</sup>We do not want to suggest any particular process of adjustment. As long as it is *monotonic* in the sense that the current reference price  $p_{\text{ref}}$  rises (falls) if it is at least  $\epsilon$  below (above) the induced equilibrium price  $p^*$  and sufficiently gradual to involve no big over-shooting, a price convention will be reached near the corresponding boundary  $\underline{p}$  or  $\bar{p}$ . Note, however, that such dynamics would have to operate on a much larger time-scale than the back-and-forth of offers considered in the baseline model. Otherwise it would be inappropriate to assume that agents behave as in a truly stationary market environment.



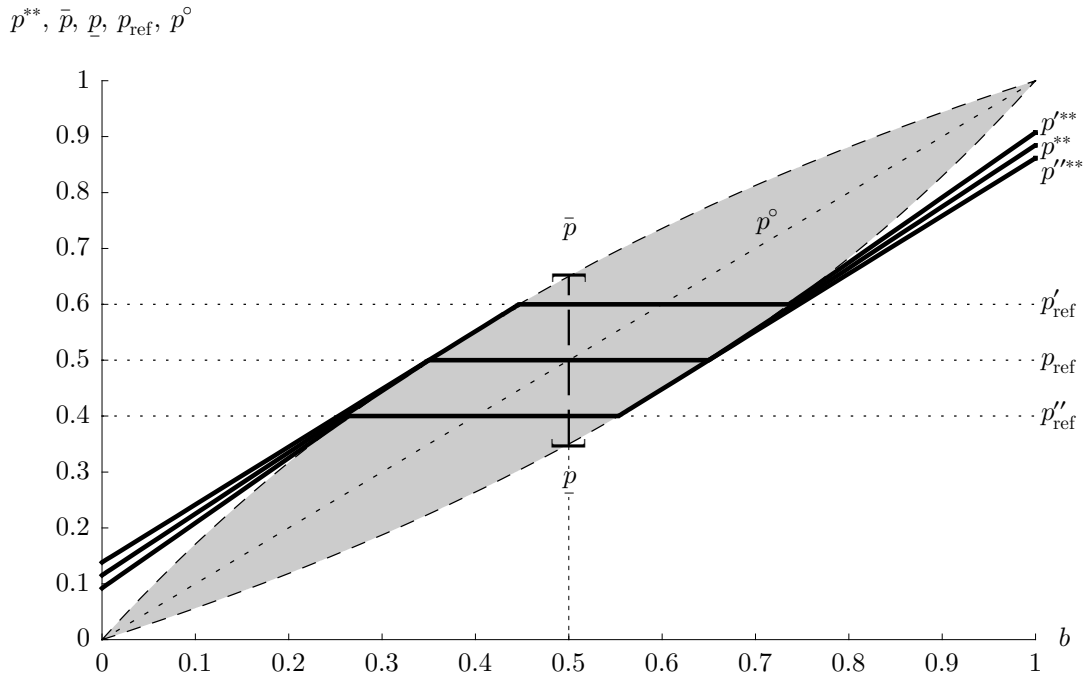


FIGURE 4.  $\epsilon$ -stable range and equilibrium prices for  $\alpha_i = \beta_i = 0.3$  and  $\delta \rightarrow 1$

This comparative static result can also be interpreted as saying that the tolerable degree of cognitive dissonance  $\epsilon$  which suffices to give rise to a *given* interval of  $\epsilon$ -stable price conventions is smaller, the larger is the discount factor  $\delta$ , i.e., the less friction is associated with rejection of an offer. In the limit of  $\delta \rightarrow 1$ , the equilibrium market price  $p^*$  is constant for the entire range of buyer-seller ratios pertaining to case (b). Then an arbitrarily small  $\epsilon > 0$  will render the entire interval  $[p, \bar{p}]$  self-confirming: any reference price  $p_{\text{ref}} \in [p, \bar{p}]$  implies an equilibrium in which the actual average price happens to be  $p_{\text{ref}}$ . This is illustrated in Figure 4. It shows the limit of equilibrium market prices for three different reference prices. For each one, there is a range of buyer shares  $b$  which would give rise to it in equilibrium. So any reference price inside the indicated grey intervals  $[p, \bar{p}]$  would be self-confirming in a robust sense: namely, for an interval of buyer shares.

Note that the difference between the prices proposed in equilibrium by a buyer and by a seller disappears as  $\delta \rightarrow 1$ . They approach

$$p^{**} \equiv \lim_{\delta \rightarrow 1} p^* = \begin{cases} b \frac{1-p_{\text{ref}}\alpha_B}{1-\alpha_B} + (1-b) \frac{p_{\text{ref}}\beta_S}{1+\beta_S} & \text{if } p_{\text{ref}} \geq \bar{p} & \text{(a)} \\ p_{\text{ref}} & \text{if } \underline{p} < p_{\text{ref}} < \bar{p} & \text{(b)} \\ b \frac{1+p_{\text{ref}}\beta_B}{1+\beta_B} - (1-b) \frac{p_{\text{ref}}\alpha_S}{1-\alpha_S} & \text{if } p_{\text{ref}} \leq \underline{p} & \text{(c)} \end{cases}$$

with

$$\bar{p} = \frac{b(1+\beta_S)}{b(1+\beta_S)+(1-b)(1-\alpha_B)}$$

and

$$\underline{p} = \frac{b(1-\alpha_S)}{b(1-\alpha_S)+(1-b)(1+\beta_B)}.$$

This would simplify to<sup>16</sup>

$$(10) \quad p^{**} = b \equiv p^\circ,$$

for “standard” preferences with  $\alpha_i = \beta_i = 0$ : the limit market price corresponds exactly to sellers’ market power as measured by the relative abundance of buyers. So the dotted 45°-line in Figure 4 illustrates how aversion to unfairly favorable deals ( $\alpha_i > 0$ ) and aversion to unfairly unfavorable deals ( $\beta_i > 0$ ) shift, e.g., a low standard equilibrium price  $p^\circ$  up towards  $p_{\text{ref}} = 0.5$  if buyers are scarce and, analogously, down towards  $p_{\text{ref}} = 0.5$  if sellers are scarce.

This qualitative observation does *not* require that both buyers and sellers are concerned about fairness, and neither that there is aversion against both unfairly favorable and unfavorable agreements. As illustrated in Figures 5(a) and (b), in order for  $p^\circ$  to be shifted downwards in a sellers’ market it is sufficient that either  $\alpha_S > 0$  or  $\beta_B > 0$ , i.e., sellers suffer from favorable deals or buyers suffer from unfavorable deals (relative to the current notion of a fair price). Similarly, Figures 5(c) and (d) illustrate that the standard equilibrium price is shifted upwards in a buyers’ market if either buyers subjectively suffer from unfairly favorable deals ( $\alpha_B > 0$ ) or sellers exhibit aversion to unfavorable deals ( $\beta_S > 0$ ).

The equilibrium price curves and ranges of self-confirming prices in Figure 4 can be understood as the superposition of the four components of Figure 5: both aversion against unfairly favorable deals by the buyer ( $\alpha_B > 0$ ) and aversion against unfairly unfavorable deals on the part of the seller ( $\beta_S > 0$ ), first, cause price stickiness in a *left* neighborhood of the buyer-seller ratio  $b = p_{\text{ref}}$  and, second, create an entire interval of prices *above* the standard equilibrium level  $p^\circ$  that would remain in place as a self-confirming endogenous reference price. Analogously, positive levels of  $\alpha_S$  (seller suffers from unduly favorable deals) and positive levels of  $\beta_B$  (buyer incurs additional subjective costs from paying more than  $p_{\text{ref}}$ ) create price stickiness in a *right* neighborhood of  $b = p_{\text{ref}}$  and, moreover, allow for  $\epsilon$ -stable reference prices below the corresponding level of  $p^\circ$ . The effects of  $\alpha_S$ ,  $\alpha_B$ ,  $\beta_S$ , and  $\beta_B$  complement one another. As can easily be checked, the magnitude of each such effect grows monotonically in the corresponding levels  $\alpha_i$  and  $\beta_i$ . For instance, there is more scope *ceteris*

<sup>16</sup>Note that equilibrium prices do not violate the tenet that agents on the short side of the market can appropriate all surplus. As first pointed out by Gale, the *flows* rather than the stocks of buyers and sellers need to be considered. Because they are equal in a stationary market, any  $p \in [0, 1]$  would be a competitive equilibrium in our setting whilst only  $p^*$  as characterized by (3)–(7) is also a strategic equilibrium. See Gale (1986a; 1986b), Rubinstein (1989), Rubinstein and Wolinsky (1990) and Kunimoto and Serrano (2004) for a more detailed discussion.

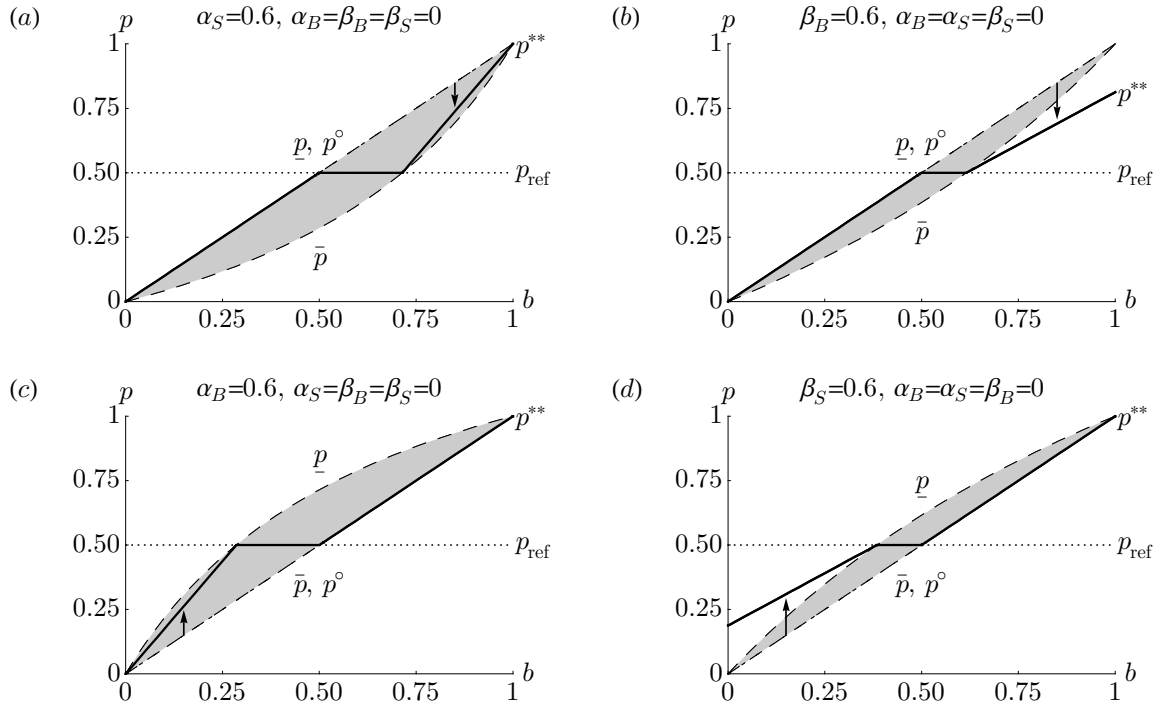


FIGURE 5. Ranges of approx. sustained reference prices for  $(\delta \rightarrow 1)$  and  $p_{\text{ref}} = 0.5$

*paribus* for a price above the standard level  $p^\circ$  to be sustained as the self-confirming fair level, the bigger are  $\alpha_B$  or  $\beta_S$ .<sup>17</sup> Or, for any given current reference price  $p_{\text{ref}}$ , the equilibrium price exhibits greater downwards stickiness in the sense of responding less when the share of buyers in the market drops from  $b = p_{\text{ref}}$ , the bigger are  $\alpha_B$  or  $\beta_S$ . The latter effect will show itself in the form of a larger interval of buyer shares  $b$  that give rise to  $p^{**} = p_{\text{ref}}$ . For instance,

$$(11) \quad p^{**} = p_{\text{ref}} = 0.5 \text{ iff } b \in \left[ \frac{1 - \alpha_B}{(1 - \alpha_B) + (1 + \beta_S)}, \frac{1 + \beta_B}{(1 - \alpha_S) + (1 + \beta_B)} \right],$$

and it is easily checked that the left endpoint of the interval decreases in  $\alpha_B$  and  $\beta_S$ , whilst the right endpoint increases in  $\alpha_S$  and  $\beta_B$ .

Returning to the stylized example of a stationary housing market, these comparative static considerations indicate that the scope for persistent price differences between two structurally identical cities A and B (in particular, involving the same reference-dependent preferences amongst the respective tenants and landlords) is bigger, the more pronounced is agents' concern with fairness. Identical average prices should only be expected to arise if fairness does not matter, i.e., when  $\alpha_i = \beta_i = 0$  for  $i \in \{B, S\}$ .

<sup>17</sup>For example,  $\frac{\partial \bar{p}}{\partial \beta_S} = \frac{(1-b)b(2-\delta)\delta(1-\alpha_B)}{((1-b)\delta(1-\alpha_B)+b(2-\delta)(1+\beta_S))^2} > 0$  for all feasible parameter constellations.

#### 4. CONCLUDING REMARKS

We have obtained three main qualitative results from our analysis of fairness in a stylized model of decentralized trade. The first, very broad finding may well be the most significant one: fairness concern can be expected to affect equilibrium market prices in a persistent way, despite the availability of complete contracts and complete information. This prediction contrasts with most of the experimental literature on behavior in markets – presumably because it has mainly looked at double auctions so far.

Our second and more specific finding is that fairness concern is likely to result in a reduced sensitivity of prices in response to variations of market fundamentals. Here, we have studied variations of the buyer-seller ratio, which is a measure of sellers' market power. We conjecture that similar results could also be obtained regarding other supply and demand shocks in a more general model, involving, for instance, heterogeneous reservation prices. Somewhat surprisingly, price sensitivity to changes in the relative numbers of buyers and sellers does not increase as the delay between the rejection of an offer and a possible re-matching vanishes. To the contrary: prices may get stickier, the *less* friction there is. The intuitive reason for this is that the bargaining advantage conferred by being on the short side of the market is linked to agents' impatience. It decreases, and eventually vanishes, as agents care less about a given time of delay. In contrast, the symmetrically distributed advantage of facing a responder who dislikes lotteries involving favorable and unfavorable deviations from the reference price is not created by discounting, and therefore does not vanish in the limit.

The third finding of our analysis concerns the possibility that reference prices are subject to psychological adaptation. If we posit, as in Ken Binmore's theory of distributional justice, that agents' notion of a fair deal will ultimately adjust if (and only if) it clashes with their actual, market power-driven bargaining experience in a pronounced way, then different fair price conventions can be associated with the same underlying market fundamentals. We have investigated the stability condition that the considered way of distributing a monetary surplus is close enough to the market outcome which would follow if this way was indeed agents' fair benchmark. Generally, many prices satisfy this requirement for a fixed market configuration. Hence, a sense of fairness which adjusts to cognitive dissonance, that possibly is implied by market fundamentals, will give rise to multiple self-confirming long-run equilibria.

The history dependence and price stickiness predicted for our very stylized stationary market can have important practical implications, and provide explanations for several real-world phenomena. In particular, fairness concern creates the potential to influence market prices by 'soft' policy measures: non-binding *recommended retail prices* published by producers possibly establish common reference points for consumers and retailers. They

can be expected to induce retail market clearing at prices nearby – and moreover to reduce the sensitivity of prices in response to demand or supply fluctuations.<sup>18</sup> Analogously, government policies such as the (lagged) publication of *rent indices* – which is common, for instance, in German municipalities – may have persistent and noticeable effects: even though landlords and tenants are generally free to deviate from the indicated price per area unit, the official figures provide a powerful benchmark; negotiators commonly point out reasons for why a lease should involve different terms. Similarly, compulsory *minimum wages* may have an economically relevant effect even after they have ceased to be a binding constraint. This has recently been documented in the laboratory by Falk et al. (2006).

Of course, our results call for more research. First, the model itself should be varied and generalized, with the goal of evaluating the robustness of our findings. Preliminary attempts to relax the presumption that all agents in the market consider the *same* reference price their benchmark suggest that the model is very robust at least to some variations. One can, for example, assume that reference prices are identical for all agents on one side of the market (namely,  $p_{\text{ref}}^i$  for  $i \in \{B, S\}$ ), but different from those on the other side (i.e.,  $p_{\text{ref}}^B \neq p_{\text{ref}}^S$ ). This gives rise to the same kind of reduced price sensitivity which we have deduced for the fully homogenous case, except that prices are now sticky around both the buyers' reference price  $p_{\text{ref}}^B$  and the sellers' reference price  $p_{\text{ref}}^S$  – corresponding to two disjoint intervals  $I^B$  and  $I^S$  of buyer-seller ratios with markedly reduced sensitivity to variations of  $b \in B^i$ . In the limit, as  $\delta \rightarrow 1$ , equilibrium prices again become perfectly flat inside  $I^B$  and  $I^S$  with a transition zone in between, where  $p^{**}$  has a slope close to unity.

Second, it would be desirable to investigate the predictions of the model empirically – in the field or, often allowing sharper targeting, by laboratory experiments. Reflecting the three main qualitative findings, a first testable hypothesis is that different reference prices – induced, for instance, by the framing of the experiment, or by initially allowing prices to settle at some level and then varying the number of buyers and sellers – produce a significant treatment effect. A second one is that the sensitivity of outcomes with respect to sellers' (or buyers') market power varies according to the three qualitative cases captured by equation (8) and, moreover, by how close the original price level was to the induced reference price. Finally, provided that stickiness around an induced benchmark can indeed be observed, how does this vary with agents' incentives to avoid delay? Does friction indeed erode the effects of fairness in bargaining markets? Investigation of these issues should require neither a market that is literally stationary, nor particularly large buyer and seller populations. In our view, a lot would already be gained by giving up the double auction paradigm for market experiments.

---

<sup>18</sup>Let us emphasize again that we have adopted a rather broad notion of fairness. We understand it as a common benchmark for the participants of a given market, as, e.g., the different local crop sharing conventions in the study of Young and Burke (2001). Consumers may not *consciously* attribute any fairness properties to the recommended retail price of an umbrella – until an opportunistic vendor starts exceeding it during heavy rain.

**Acknowledgements.** We thank N. Shunda for helpful comments on an earlier draft, and acknowledge constructive discussions with participants of the 3<sup>rd</sup> World Congress of the Game Theory Society, the 2007 Nürnberg BGPE conference, and seminars at Berlin, Hamburg and Maastricht. Napel acknowledges the hospitality of the Institute for Economic Development at Boston University.

## REFERENCES

- Akerlof, G. A. and W. T. Dickens (1982). The economic consequences of cognitive dissonance. *American Economic Review* 72(3), 307–319.
- Albrecht, J., A. Anderson, E. Smith, and S. Vroman (2007). Opportunistic matching in the housing market. *International Economic Review* 48(2), 641–664.
- Ball, L. and N. G. Mankiw (1994). A sticky-price manifesto. *Carnegie-Rochester Conference Series on Public Policy* 41, 127–151.
- Binmore, K., P. Morgan, A. Shaked, and J. Sutton (1991). Do people exploit their bargaining power? An experimental study. *Games and Economic Behavior* 3(3), 295–322.
- Binmore, K., J. Swierzbinski, S. Hsu, and C. Proulx (1993). Focal points and bargaining. *International Journal of Game Theory* 22(4), 381–409.
- Binmore, K. G. (1994). *Game Theory and the Social Contract – Volume I: Playing Fair*. Cambridge, MA.
- Binmore, K. G. (1998). *Game Theory and the Social Contract – Volume II: Just Playing*. Cambridge, MA.
- Binmore, K. G. and M. J. Herrero (1988). Matching and bargaining in dynamic markets. *Review of Economic Studies* LV, 17–31.
- Bolton, G. and A. Ockenfels (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review* 90 (1), 166–193.
- Bolton, L. E., L. Warlop, and J. W. Alba (2003). Consumer perceptions of price (un)fairness. *Journal of Consumer Research* 29, 474–491.
- Brown, M., A. Falk, and E. Fehr (2004). Relational contracts and the nature of market interactions. *Econometrica* 72(3), 747–780.
- Camerer, C. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press.
- Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117(3), 817–869.
- Cooper, J. (2007). *Cognitive Dissonance: 50 Years of a Classic Theory*. Sage Publications Ltd.
- Diamond, P. A. and E. Maskin (1979). An equilibrium analysis of search and breach of contract, I: Steady states. *Bell Journal of Economics* 10(1), 282–316.
- Dufwenberg, M. and G. Kirchsteiger (2004). A theory of sequential reciprocity. *Games and Economic Behavior* 47(2), 268–298.

- Ewerhart, C. (2004). The effect of sunk costs on the outcome of alternating-offers bargaining between inequity-averse agents. Institute for Empirical Research in Economics, University of Zurich, Working Paper Series, Working Paper No. 203.
- Falk, A., E. Fehr, and C. Zehnder (2006). Fairness perceptions and reservation wages - the behavioral effects of minimum wage laws. *Quarterly Journal of Economics* 121(4), 1347–1381.
- Falk, A. and U. Fischbacher (2006). A theory of reciprocity. *Games and Economic Behavior* 54(2), 293–315.
- Fehr, E. and K. Schmidt (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics* 114, 817–868.
- Fehr, E. and J.-R. Tyran (2001). Does money illusion matter? *American Economic Review* 91(5), 1239–1262.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Franciosi, R., P. Kujal, R. Michelitsch, V. Smith, and G. Deng (1995). Fairness: Effect on temporary and equilibrium prices in posted-offer markets. *Economic Journal* 105(431), 938–950.
- Gale, D. (1986a). Bargaining and competition Part I: Characterization. *Econometrica* 54(4), 785–806.
- Gale, D. (1986b). Bargaining and competition Part II: Existence. *Econometrica* 54(4), 807–818.
- Gode, D. K. and S. Sunder (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy* 101(1), 119–137.
- Golosov, M. and R. E. Lucas (2007). Menu costs and Phillips curves. *Journal of Political Economy* 115, 171–199.
- Kachelmeier, S. J., S. T. Limberg, and M. S. Schadewald (1991). Fairness in markets: A laboratory investigation. *Journal of Economic Psychology* 12(3), 447–464.
- Kahneman, D., J. L. Knetsch, and R. H. Thaler (1986). Fairness and the assumptions of economics. *Journal of Business* 59(4), 285–300.
- Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review* 90(4), 1072–1091.
- Köszegi, B. and M. Rabin (2006). A model of reference dependent preferences. *Quarterly Journal of Economics* 121(4), 1133–1165.
- Krainer, J. and S. F. LeRoy (2002). Equilibrium valuation of illiquid assets. *Economic Theory* 19(2), 223–242.
- Kunimoto, T. and R. Serrano (2004). Bargaining and competition revisited. *Journal of Economic Theory* 115(1), 78–88.

- Mortensen, D. (1982). The matching process as a noncooperative bargaining game. In J. J. McCall (Ed.), *The Economics of Information and Uncertainty*, pp. 233–254. The University of Chicago Press.
- Mortensen, D. T. and C. A. Pissarides (1999). New developments in models of search in the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, Chapter 39, pp. 2567–2627. Elsevier.
- Nail, P. R., J. E. Misak, and R. M. Davis (2004). Self-affirmation versus self-consistency: a comparison of two competing self-theories of dissonance phenomena. *Personality and Individual Differences* 36, 1893–1905.
- Rocheteau, G. and R. Wright (2005). Money in search equilibrium, in competitive equilibrium, and in competitive search equilibrium. *Econometrica* 73(1), 175–202.
- Rogerson, R., R. Shimer, and R. Wright (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature* 43(4), 959–988.
- Rubinstein, A. (1989). Competitive equilibrium in a market with decentralized trade and strategic behavior: An introduction. In G. R. Feiwel (Ed.), *The Economics of Imperfect Competition and Employment: Joan Robinson and Beyond*, Chapter 7, pp. 243–259. Basingstoke: Macmillan.
- Rubinstein, A. and A. Wolinsky (1985). Equilibrium in a market with sequential bargaining. *Econometrica* 53, 1133–1150.
- Rubinstein, A. and A. Wolinsky (1990). Decentralized trading, strategic behaviour and the walrasian outcome. *Review of Economic Studies* 57, 63–78.
- Shi, S. (2001). Liquidity, bargaining, and multiple equilibria in a search monetary model. *Annals of Economics and Finance* 2, 325–251.
- Taylor, J. B. (1999). Staggered price and wage setting in macroeconomics. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1 of *Handbook of Macroeconomics*, Chapter 15, pp. 1009–1050. Elsevier.
- Woodford, M. (2002). Imperfect common knowledge and the effects of monetary policy. In P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford (Eds.), *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*. Princeton University Press.
- Xia, L., K. B. Monroe, and J. L. Cox (2004). The price is unfair! A conceptual framework of price fairness perceptions. *Journal of Marketing* 68, 1–15.
- Young, H. P. and M. A. Burke (2001). Competition and custom in economic contracts: A case study of Illinois agriculture. *American Economic Review* 91(3), 559–573.