

Learning to Like What You Have – Explaining the Endowment Effect*–

Steffen Huck
Department of Economics
Humboldt University, Berlin

Georg Kirchsteiger
CentER for Economic Research
Tilburg University

Jörg Oechssler[†]
Department of Economics
Humboldt University, Berlin

April 15, 1997

Abstract

The *endowment effect* describes the fact that people demand much more to give up an object than they are willing to spend to acquire it. The existence of this effect has been documented in numerous experiments. We attempt to explain this effect by showing that evolution favors individuals whose preferences embody an endowment effect. The reason is that an endowment effect improves one's bargaining position in bilateral trades. We show that for a general class of evolutionary processes almost all individuals will have a strictly positive and finite endowment effect.

JEL- classification numbers: C73, C79, D00.

*We thank Ted Bergstrom, Lorne Carmichael, Simon Gächter, Andras Löffler, Clemens Puppe and Frank Riedel for helpful comments.

[†]Institut für Wirtschaftstheorie, Spandauer Str. 1, 10178 Berlin, Germany,
FAX: +49-30-2093-5619, E-mail: oechssler@wiwi.hu-berlin.de

1 Introduction

The endowment effect describes the fact that people demand much more to give up an object than they are willing to spend to acquire it (Thaler, 1980). This phenomenon that people attach higher values to goods if they are in their possession is very well established in the experimental literature (see Kahneman *et al.*, 1991, for an overview). In a famous experiment Kahneman *et al.* (1990) distributed coffee mugs to every other student in a classroom. When asked to state their valuations for the mugs, the ‘mug-owners’ had on average a much higher valuation than the other students. In fact, the willingness to accept (WTA) by the mug-owners was about twice as high as the willingness to pay (WTP) by the remaining students. Such extreme differences seem unreasonable but they have been confirmed in countless other experiments (see e.g. Knetsch, 1989, or Knetsch and Sinden, 1984) and must by now be accepted as a stylized fact.¹

The endowment effect is not just an experimental curiosity. As pointed out by Kahneman *et al.* (1990), the endowment effect questions the validity of the Coase Theorem. The Coase Theorem states that the allocation of property rights does not influence the way external effects are internalized by the market. However, if the endowment effect influences the valuations, property rights do matter. One real life example where the endowment effect might play an important role is the market for so-called “reverse mortgages” (see e.g., Weinrobe, 1988). Reverse mortgages are contracts in which a home owner sells back his property to the bank in exchange for an annuity often including a life insurance component. Those contracts seem to be sensible instruments if one is risk averse and wants to smooth consumption over the life cycle. Although reverse mortgages are available in the US since 1981, there has been little demand for them (Venti and Wise, 1990). The endowment effect may present an explanation for this lack of demand since it would lead home owners to attach excessively high values to their property.

Knez and Smith (1987) and Coursey *et al.* (1987) challenged the view that the experiments are proof for preferences that depend on endowments. Coursey *et al.* found that the differences between WTA and WTP are smaller when a Vickrey auction is used to determine the valuations or when subjects have the opportunity to gain experience in a market setting.²

¹Kahneman *et al.* (1990) alone report experiments with more than 700 subjects.

²However, substantial differences between WTA and WTP remain also in the Coursey *et al.* (1987) experiment. For a critique of the Coursey *et al.* experiment see Knetsch and Sinden (1987).

They argue that the observed behavior may be due to the fact that subjects mistakenly apply bargaining behavior which is sensible in normal bargaining situations but inappropriate in the experiments, namely to understate one's WTP and to overstate one's WTA. Coursey *et al.* (1987) conclude that the endowment effect should play a lesser role in market environments or auctions. In reaction to this Kahneman *et al.* (1990) used markets in their experiments and still found a very strong endowment effect. It seems that the observed endowment effect cannot be explained by strategic considerations. Rather, it truly reflects endowment dependent preferences.

Preferences of an individual are determined by his genes and/or his socialization. In both cases they can be regarded as the product of an evolutionary process, either biological or cultural. Therefore, if one wants to explain the existence of the endowment effect, one has to ask for the evolutionary benefits of having such preferences. The purpose of our paper is to analyze theoretically whether people are better off with an endowment effect than without. Furthermore, we ask whether plausible evolutionary dynamics lead to populations which exhibit an endowment effect.

In our model we consider subjects endowed with different goods. Initially, one group is endowed only with good x , the “ x -owners”, while the other group is endowed only with y , the “ y -owners”. As an example consider a population made up of hunters and farmers. It is probably healthy to consume a balanced diet of meat and cereal, which is only possible if hunters and farmers engage in trade.

During most of the history of mankind trade took place within small groups that did not resemble a market. Furthermore, as children and teenagers — when our socialization takes place — we trade things with friends or siblings. Again, these trades take place not in market situations but rather in bilateral or multilateral bargaining situations. Hence, genetically determined preferences as well as those determined by socialization should mainly be formed in individual bargaining situations rather than in market situations.³ Therefore, we have chosen a model in which the terms

³There exists a lot of experimental evidence that market clearing equilibria arise even in small groups. In these experiments, however, trade takes place on centralized markets with well specified rules like the oral double auction. Such markets are neither typical for most of the human history nor for trades conducted in early periods of individual life. Therefore, decentralised markets without specified rules like that tested by Chamberlain (1948) seem to be more appropriate for the examination of the problem at hand. We have no indication that such markets clear. Rather, it seems that on such markets homogenous goods are traded at different prices. As a simplification of such a situation we have chosen a bilateral bargaining model which is, of course, an extreme example of a trading institution

of trade are determined by bargaining and not by a market mechanism.

It is important to differentiate two concepts, the (objective) *fitness* an individual derives from the consumption of the two goods and the (subjective) *utility* derived from those. In the hunter–farmer example above fitness is given by the “nutritional value” of the diet, with balanced diets yielding higher fitness. Utility may differ from this when there is an endowment effect. A hunter with an endowment effect would like meat more than is good for him in terms of fitness. Why would a hunter develop an endowment effect, which lets him consume meat and cereal in suboptimal proportions? The answer is that an endowment effect may improve his bargaining position and result in higher fitness despite the distortion from the optimal rate of substitution between meat and cereal.

We assume that the bargaining takes place according to the Nash bargaining solution. Developing an endowment effect in this context has two effects. First, an endowment effect distorts the marginal rate of substitution, at which a person is willing to trade, away from the “objective” marginal rate of substitution given by the fitness function. But secondly, the endowment effect moves the threat point in one’s favor.

Since we are interested in the evolutionary success of the endowment effect, we have to compare the fitness of individuals with and without an endowment effect. The endowment effect is formalized by an “endowment parameter” which is an additive term in the utility function and can take any positive real value. We make the straightforward assumption that the share of individuals with a given endowment parameter grows faster than the share of other individuals if and only if the former receive a higher average payoff (measured in fitness) than the latter. These so-called growth monotonic dynamics (Weibull, 1995) can be seen as the result of genetic evolution or as the result of (cultural) imitation.

Our main result is that it always pays to develop an endowment effect. We show that starting from any interior initial condition the evolutionary process will drive down the share of individuals without an endowment parameter to zero. Preferences that exhibit an endowment effect can therefore be explained by the success people endowed with these preferences had in the past.

Notice the difference of our explanation to that of Coursey *et al.* (1987). We do not assume that individuals misrepresent their true valuations for

that allows for different prices. We think, however, that as long as the “law of one price” does not hold our results remain valid.

strategic reasons. Rather we assume that people behave in accordance with their preferences, and it turns out that people with endowment preferences do evolutionary better than those without.⁴

There are several papers in the recent literature that study bargaining behavior from an evolutionary viewpoint, e.g. Young (1993), Gale *et al.* (1995), Huck and Oechssler (1996), Ellingsen (1996) and Carmichael and MacLeod (1996). The last three papers are closest to the current paper. Huck and Oechssler (1996) study the evolution of preferences in the Ultimatum Game using the indirect evolutionary approach, which is also used in the current paper. Ellingsen's (1996) paper is concerned with the evolution of bargaining behavior in the Nash demand game. In his setup there are two types of players, "rational" (or responsive) types and "obstinate" types. Ellingsen shows that only a mix of responsive types and obstinate types who demand an even split can be evolutionary stable.

In independent work Carmichael and MacLeod (1996) arrive at similar conclusions as the current paper. They show for an example with a square-root fitness function that it may be advantageous to develop an endowment effect in bargaining situations. However, they do not use (static or dynamic) evolutionary concepts. Rather, they look for efficient Nash equilibria, which are, as the recent literature on evolutionary games shows (see e.g. Kandori, Mailath and Rob, 1993), not always the equilibria selected by evolution.

The remainder of the paper is organized as follows. In the next section we describe the basics of the model. Then we analyze the bargaining process. In Section 4 we show that the endowment effect is an evolutionary advantage and in Section 5 we analyze the dynamics of the evolutionary process. In the last section conclusions are drawn.

2 The model

Consider an economy with two goods, x and y . There are two types of individuals, those who have an endowment of x only, the " x -owners", and those with an endowment of y , the " y -owners". Each individual has an endowment of one unit of his good. We suppose that there is a continuum of individuals of each type with the relative size of the two populations being constant. Individuals from both populations are randomly matched to engage in bilateral trade.

⁴This "indirect evolutionary approach" was developed by Güth and Yaari (1992).

Individuals derive fitness from the consumption of x and y according to the objective fitness function $\bar{F}(x, y) : [0, 1]^2 \rightarrow \mathbb{R}$, which is the same for all individuals. We assume that $\bar{F}(x, y)$ is strictly increasing in both arguments, strictly concave, twice continuously differentiable and bounded. The partial derivatives $\bar{F}_x(x, y)$ and $\bar{F}_y(x, y)$ are bounded and the cross-derivative $\bar{F}_{xy}(x, y)$ is strictly positive. Furthermore, we impose a weak form of symmetry between x and y in the sense that their contribution to fitness is not too different. Namely, we assume that consuming all of one good is better than consuming only half of the other, i.e., $\bar{F}(1, 0) > \bar{F}(0, \frac{1}{2})$ and $\bar{F}(0, 1) > \bar{F}(\frac{1}{2}, 0)$.

Individuals also derive subjective utility from the consumption of x and y , which may differ from objective fitness. In particular, we assume that x -owners may develop a utility function of the form

$$\bar{U}_1(x_1, y_1) := \bar{F}(x_1, y_1) + e_1 x_1.$$

Similarly for y -owners

$$\bar{U}_2(x_2, y_2) := \bar{F}(x_2, y_2) + e_2 y_2.$$

The additional term with the *endowment parameter* $e_i \geq 0$ signifies an increased preference for the good one owns. For computational simplicity we restrict our attention to this class of preferences. But notice that the main result of our paper, which relates to the existence of an endowment effect, does not depend on this restriction. Clearly, it may be the case that we exclude preferences that do better than those allowed in our model. But these better preferences cannot be preferences without an endowment effect ($e_i = 0$), since such preferences are included in the class of preferences we allow for.

For notational convenience we will from now on work with incremental fitness, which is the difference between $\bar{F}(x, y)$ and the fitness an individual receives from consuming his endowment. For x -owners incremental fitness is

$$F(x, y) := \bar{F}(x, y) - \bar{F}(1, 0)$$

and similarly for y -owners

$$F_2(x, y) := \bar{F}(1 - x, 1 - y) - \bar{F}(0, 1),$$

where we generally drop the subscript “1”, i.e. $x = x_1, F(x, y) = F_1(x, y)$ etc. Furthermore, since total endowment of both goods is one, we can replace

x_2 by $1 - x$. Partial derivatives of the y -owner are denoted by $F_{2z}(x, y) := F_z(1 - x, 1 - y)$ for $z = x, y$.

Since affine transformations of the utility function do not affect the analysis, we normalize the utility functions in the following way.

$$U(x, y) := F(x, y) + e_1x$$

$$U_2(x, y) := F_2(x, y) + e_2(1 - y).$$

Note that with this transformation $U(0, 0) < 0 \leq e_1 = U(1, 0) < U(1, 1)$.

3 The bargaining process

Consider now a bargaining situations in which two individuals, one from each type, are randomly matched to bargain about x and y . Let us denote the feasible set of allocations in utility space by S . Since $\overline{F}(x, y)$ is bounded and strictly concave, S satisfies all standard assumptions, in particular, for finite endowment parameters, e_1 and e_2 , S is bounded, closed and strictly convex. Due to our normalization the threat point d is simply given by (e_1, e_2) .

We assume that if Pareto improving allocations are feasible, the Nash bargaining solution is being implemented.⁵ As usual, the Nash bargaining solution is obtained by maximizing the Nash product $N(x, y)$

$$(x^*, y^*) := \arg \max_{(x, y) \in [0, 1]^2} N(x, y), \quad (1)$$

where

$$N(x, y) := (U(x, y) - e_1)(U_2(x, y) - e_2) = (F(x, y) + e_1x - e_1)(F_2(x, y) - e_2y).$$

If no Pareto improving allocation is feasible, individuals simply consume their endowments.

Lemma 1 *If there exist Pareto improving allocations, then there is an interior Nash solution, $(x^*, y^*) \in (0, 1)^2$. Furthermore, this solution is unique.*

⁵In principle, this implies that the utility functions of both individuals in a match are commonly known. Ellingsen (1996) provides strong arguments in favor of this assumption.

Proof If Pareto improving allocations exist, the maximized Nash product $N(x^*, y^*)$ must be strictly positive, i.e. both individuals must be strictly better off than in their endowment.

Suppose (x^*, y^*) were on a boundary. There are four cases.

(1) $y^* = 0, x^* \in [0, 1]$. Obviously, the x -owner is (weakly) worse off than with his endowment since

$$U(x^*, 0) = F(x^*, 0) + e_1 x^* \leq F(1, 0) + e_1 = U(1, 0),$$

which yields a contradiction.

(2) $x^* = 1, y^* \in (0, 1]$. Similar to case (1) with the y -owner being worse off.

(3) $x^* = 0, y^* \in (0, 1]$. For this to be a Nash solution it must hold that

$$\begin{aligned} N_x(x^*, y^*) &= (F_x + e_1)(F_2 - e_2 y^*) - F_{2x}(F + e_1 x^* - e_1) \leq 0 \\ N_y(x^*, y^*) &= F_y(F_2 - e_2 y^*) - (F_{2y} + e_2)(F + e_1 x^* - e_1) \geq 0 \end{aligned}$$

which implies that

$$MRS_1 = \frac{F_x + e_1}{F_y} \leq \frac{F_{2x}}{F_{2y} + e_2} = MRS_2. \quad (2)$$

For the x -owner to be better off than with his endowment, we must have that $y^* \geq \frac{1}{2}$. Because of $F_{xy} > 0$ and $F_{xx} < 0$ it follows that

$$F_x(0, y^*) > F_x(1, 1 - y^*) = F_{2x}(0, y^*)$$

and

$$F_y(0, y^*) < F_y(1, 1 - y^*) = F_{2y}(0, y^*),$$

but this contradicts (2).

(4) $x^* \in (0, 1), y^* = 1$. This case is equivalent to case (3).

To prove uniqueness, suppose there exist two interior Nash solutions (x, y) and (x', y') . Since utility functions are concave both individuals would prefer any convex combination of (x, y) and (x', y') over (x, y) and (x', y') . These convex combinations are feasible which yields a contradiction. ■

Lemma 2 *A Pareto improving bargaining solution exists iff*

$$F_x(0, 1)F_y(1, 0) - F_x(1, 0)F_y(0, 1) > e_1 F_x(0, 1) + e_2 F_x(1, 0) + e_1 e_2.$$

Proof A Pareto improving bargaining solution exists iff the problem

$$\max_{x,y} U_2(x,y) \quad \text{s.t.} \quad F(x,y) + e_1 x = e_1$$

has a value $U_2 > e_2$. Let $\tilde{x}(y)$ denote the x that solves the constraint for a given y . Implicitly differentiating we find that

$$\frac{d\tilde{x}}{dy} = \frac{-F_y}{F_x + e_1} < 0$$

and

$$\frac{d^2\tilde{x}}{dy^2} = \frac{2(F_x + e_1)F_y F_{xy} - (F_y)^2 F_{xx} - (F_x)^2 F_{yy}}{(F_y)^3} > 0.$$

Substituting $\tilde{x}(y)$ into $U_2(x,y)$ and differentiating yields

$$\frac{dU_2}{dy} = -F_{2x} \frac{d\tilde{x}}{dy} - F_{2y} - e_2$$

and

$$\frac{d^2U_2}{dy^2} = \left(\frac{d\tilde{x}}{dy}\right)^2 F_{2xx} + \frac{d\tilde{x}}{dy} F_{2xy} - \frac{d^2\tilde{x}}{dy^2} F_{2x} + F_{2yy} < 0.$$

Since $U_2(y)$ is concave, a Pareto improving allocation exists iff

$$\left. \frac{dU_2}{dy} \right|_{y=0} > 0$$

or

$$F_{2x}(1,0) \frac{F_y(1,0)}{F_x(1,0) + e_1} - F_{2y}(1,0) - e_2 > 0.$$

Noting that $F_{2x}(1,0) = F_x(0,1)$ etc. one gets the above condition. ■

That is, for a given e_2 there is a critical upper bound for e_1 (and equivalently for e_2) given by

$$\bar{e}_1(e_2) = \frac{F_x(0,1)F_y(1,0)}{F_y(0,1) + e_2} - F_x(1,0). \quad (3)$$

If either $e_1 \geq \bar{e}_1(e_2)$ or, equivalently, $e_2 \geq \bar{e}_2(e_1)$, no Pareto improving allocation exists and no trade will take place. Note that $\bar{e}_1(e_2)$ is continuous and strictly decreasing in e_2 . The maximal value of $\bar{e}_i(e_j)$ is therefore $\hat{e}_i := \bar{e}_i(0)$. Analogously, one obtains $\bar{e}_i(\hat{e}_j) = 0$.

4 Does it pay to develop an endowment effect?

Clearly, the outcome of the Nash bargaining solution depends on the endowment parameters e_1 and e_2 . There are two effects: On the one hand, developing an endowment effect changes the threat point in one's favor. But on the other hand, the endowment effect distorts the consumption mix away from the optimal (in terms of fitness) mix since the marginal rates of substitution in terms of fitness and in terms of utility become different.

Now compare an x -owner with endowment parameter $e_1 = 0$ and an x -owner with a slightly positive endowment parameter both being matched with a y -owner with $e_2 < \hat{e}_2$. Which of the x -owners would be better off in these encounters? The next proposition shows that, in fact, it always pays to develop at least a small endowment effect. Individuals with a small endowment effect earn more in terms of fitness than others without an endowment effect. This is due to the fact that the positive effect caused by the improvement of the bargaining position overcompensates the negative effect caused by the distortion of the marginal rate of substitution.

Proposition 1 *The fitness of an individual without an endowment effect ($e_i = 0$) is strictly increasing in the endowment parameter e_i given any endowment parameter of the other player $e_j < \hat{e}_j$.*

Note that for values $e_j \geq \hat{e}_j$ no Pareto improving allocation exists as shown in Lemma 2. Therefore, developing an endowment effect in these cases does not help (but does not harm either).

Proof We will show that for all $e_2 < \hat{e}_2$

$$\frac{\partial F(x^*(e_1, e_2), y^*(e_1, e_2))}{\partial e_1} = F_x \frac{\partial x^*}{\partial e_1} + F_y \frac{\partial y^*}{\partial e_1} > 0 \quad (4)$$

at $e_1 = 0$. Differentiating the first order conditions for the problem (1), $N_x = 0$ and $N_y = 0$, with respect to e_1 , we get that

$$\frac{\partial x^*}{\partial e_1} = \frac{N_{ye_1} N_{xy} - N_{xe_1} N_{yy}}{N_{xx} N_{yy} - (N_{xy})^2} \text{ and } \frac{\partial y^*}{\partial e_1} = \frac{N_{xe_1} N_{xy} - N_{ye_1} N_{xx}}{N_{xx} N_{yy} - (N_{xy})^2}. \quad (5)$$

As shown in Lemma 1 the solution to (1) is an interior maximum if a Pareto improving allocation is feasible. By Lemma 2 such an allocation is feasible for all $e_2 < \hat{e}_2$. Therefore, the second order condition $N_{xx} N_{yy} - (N_{xy})^2 \geq 0$ is satisfied. What remains to show is that at $e_1 = 0$

$$N_{xe_1} (F_y N_{xy} - F_x N_{yy}) + N_{ye_1} (F_x N_{xy} - F_y N_{xx}) > 0.$$

Both, N_{xe_1} and N_{ye_1} are positive,

$$\begin{aligned} N_{xe_1} &= F_2 - e_2y + F_{2x}(1-x) > 0 \\ N_{ye_1} &= (1-x)(F_{2y} + e_2) \geq 0. \end{aligned}$$

Note that $F_2 > e_2y$. Otherwise the allocation would not be Pareto improving.

Thus it suffices to show that

$$F_y N_{xy} - F_x N_{yy} > 0 \quad (6)$$

and

$$F_x N_{xy} - F_y N_{xx} > 0. \quad (7)$$

At an interior solution $MRS_1 = MRS_2$, which implies that

$$\frac{F_x + e_1}{F_y} = \frac{F_{2x}}{F_{2y} + e_2}.$$

Thus, for $e_1 = 0$, we have

$$F_y F_{2x} = F_x F_{2y} + F_x e_2. \quad (8)$$

Given that

$$N_{xy} = F_{xy}(F_2 - e_2y) + F_{2xy}(F + e_1x - e_1) - F_x(e_2 + F_{2y}) - F_y F_{2x} - e_1 e_2 - e_1 F_{2y}$$

we can define

$$N' := -2F_y F_{2x} < N_{xy},$$

where the inequality follows from (8). To verify claim (6) we show that $F_y N' - F_x N_{yy} > 0$ at $e_1 = 0$.

$$\begin{aligned} &F_y N' - F_x N_{yy} \\ &= 2F_x F_y e_2 + 2F_y (F_x F_{2y} - F_y F_{2x}) \underbrace{- F_x F_{yy} (F_2 - e_2y)}_{>0} \underbrace{- F_x F_{2yy} F}_{>0}. \end{aligned}$$

By (8) it follows that

$$2F_x F_y e_2 + 2F_y (F_x F_{2y} - F_y F_{2x}) = 0,$$

which proves the first claim.

The second claim is that $F_x N' - F_y N_{xx} > 0$. But this follows immediately:

$$\begin{aligned}
& F_x N' - F_y N_{xx} \\
= & -2F_x F_y F_{2x} - F_y (F_{2xx} F - 2F_x F_{2x} + F_{xx} (F_2 - e_2 y)) \\
= & \underbrace{-F_y F_{2xx} F}_{>0} - \underbrace{F_y F_{xx} (F_2 + e_2 y)}_{>0}. \blacksquare
\end{aligned}$$

Defining $H(e_j) = \{e_i^* : e_i^* = \arg \max_{e_i} F(x^*(e_1, e_2), y^*(e_1, e_2))\}$ one can restate Proposition 1 as

$$e_j < \hat{e}_j \Rightarrow \inf H(e_j) > 0.$$

Due to the Theorem of the Maximum (Berge, 1963) the correspondence $H(\cdot)$ is upper hemi-continuous. It is important to note that

$$\lim_{e_j \rightarrow \hat{e}_j} \sup H(e_j) = 0.$$

The latter follows since only for $e_i \in [0, \bar{e}_i(e_j))$ gains of trade are feasible. As e_j converges to \hat{e}_j the upper bound of this interval converges to 0.

Proposition 1 is concerned with the value of having an endowment effect vis à vis a trading partner with a specific given endowment parameter. In our model, however, the partner is chosen randomly from a population, which is, in general, not monomorphic with respect to the endowment parameter. Hence, one has to analyze the expected value of an endowment effect vis à vis a distribution of different partners, i.e. a distribution of endowment parameters.

Recall our assumption that there are two infinite populations of x - and y -owners. Let $(\mathbb{R}_+, \mathcal{F}, \mu_i)$ denote a probability space, where the probability measure μ_i specifies the distribution of endowment parameters (“types”) in population i and \mathcal{F} is the Borel algebra. The support of μ_j is denoted by $\Sigma_j \subset \mathbb{R}_+$. Finally, the distance between two probability measures μ_i and μ'_i is measured by the sup-norm (variational norm, see e.g. Bomze and Pötscher, 1989, p. 47)

$$\|\mu_i - \mu'_i\| := \sup_{A \in \mathcal{F}} |\mu_i(A) - \mu'_i(A)|.$$

A population game is defined by the pair of probability measures μ_1, μ_2 , and by a pair of average fitness functions

$$\begin{aligned}
\tilde{F}_1(e_1, \mu_2) &= \int_0^\infty F(x^*(e_1, e_2), y^*(e_1, e_2)) d\mu_2(e_2) \\
\tilde{F}_2(\mu_1, e_2) &= \int_0^\infty F_2(x^*(e_1, e_2), y^*(e_1, e_2)) d\mu_1(e_1)
\end{aligned}$$

which specifies the average fitness a type e_i receives in population i given the distribution of types in population j .

Corollary 1 *The fitness of a player without an endowment effect is strictly increasing in his endowment parameter e_i given any distribution of types in the other population with $\mu_j([0, \widehat{e}_j]) > 0$.*

Proof Since (4) holds for any $e_j < \widehat{e}_j$, average fitness $\widetilde{F}_i(e_i, \mu_j)$ is strictly increasing in e_i at $e_i = 0$ for any distribution with $\mu_j([0, \widehat{e}_j]) > 0$. Since

$$\frac{\partial F_i(x^*(e_1, e_2), y^*(e_1, e_2))}{\partial e_i} = 0$$

for all $e_j \geq \widehat{e}_j$, the corollary follows. ■

5 Evolutionary dynamics

Proposition 1 and Corollary 1 show that preferences embodying an endowment effect are evolutionarily advantageous. In this section we will introduce a plausible class of evolutionary dynamics in order to investigate whether this advantage is strong enough to lead to the extinction of “rational” individuals, that is, individuals without an endowment effect. Furthermore, if there is an advantage to having an endowment effect, the question arises whether it is so strong that individuals will engage in an “arm’s race” ending up with infinite endowment parameters, which, of course, would result in an inefficient no-trade situation. Finally, we ask under what conditions stable monomorphic populations exist.

The evolutionary dynamics describe how the distribution of types in the two populations changes over time. Let Ω denote the state space of the process, i.e. Ω is the set of all probability measures over endowment parameters in both populations. Let $\omega := (\mu_1, \mu_2)$ be a typical element of Ω . The flow of the system starting from an initial state ω^0 is denoted by $\varphi^t(\omega^0)$.

We will assume that the evolutionary dynamics are regular⁶ (Weibull, 1995) and payoff monotonic. The latter assumption simply says that types with higher average fitness have higher growth rates, or formally

⁶A process φ is regular if it is Lipschitz continuous and if it remains in the simplex at all times, i.e. $\dot{\mu}(\Omega) = 0$.

Definition 1 A dynamic process $\{\varphi^t(\omega)\}$ is called payoff monotonic if for $i = 1, 2$ and for all A and $A' \in \mathcal{F}$

$$\frac{1}{\dot{\mu}_i^t(A)} \int_A \tilde{F}(e_i, \mu_j) d\mu_i^t(e_i) > \frac{1}{\dot{\mu}_i^t(A')} \int_{A'} \tilde{F}(e_i, \mu_j) d\mu_i^t(e_i) \Leftrightarrow \frac{\dot{\mu}_i^t(A)}{\mu_i^t(A)} > \frac{\dot{\mu}_i^t(A')}{\mu_i^t(A')},$$

where $\dot{\mu}_i^t(A)$ denotes the time derivative of $\mu_i^t(A)$.

An example for the class of processes which satisfy the above assumptions is the continuous version of the replicator dynamics, which are defined as ⁷

$$\dot{\mu}_i(A) = \int_A \tilde{F}(e_i, \mu_j) d\mu_i(e_i) - \mu_i(A) \int_0^\infty \tilde{F}(e_i, \mu_j) d\mu_i(e_i).$$

Most plausible evolutionary or learning processes will have to satisfy payoff monotonicity. In particular, payoff monotonic dynamics describe well cultural imitation processes (see Weibull, 1995).

A state ω is called stationary if $\varphi^t(\omega) = \omega$ for all t . Trivially, all states are stationary in which the cardinalities of both, Σ_1 and Σ_2 , are one, i.e. if both populations are monomorphic. Additionally, there may also be absorbing states with polymorphic populations. Given the multiplicity of stationary states we are interested in their stability properties. A stationary state ω is (Lyapunov) *stable* if for every neighborhood U of ω there exists a neighborhood U' of ω in U such that $\varphi^t(\omega^0) \in U$ for all $\omega^0 \in U'$ and $t \geq 0$.

We will demonstrate first that all states in which some individuals have no endowment effect or states in which some individuals have an excessively high (or even infinite) endowment effect are not stable. We prove this by showing that no “boundary” state can be (Lyapunov) stable. Boundary states are states in which in one of the populations there is either a positive probability mass of individuals with no endowment effect or with exceedingly high endowment effects. That is, $\omega \in \Omega$ is a *boundary state* if in at least one population i either $\mu_i(\{0\}) > 0$, or $\mu_i([\hat{e}_i, \infty]) > 0$. Let B denote the set of all boundary states and \overline{B} all stationary points in B .

Proposition 2 All $\omega \in \overline{B}$ are unstable.

The proof of the proposition is based on the following lemma. Let $\delta_i(\mu_j) := \inf\{e_i : e_i = \arg \max \tilde{F}_i(\cdot, \mu_j)\}$ and $\delta_i^*(\omega) := \inf\{\delta_i(\mu_j) : \exists t > 0$

⁷Usually, the replicator dynamics are defined for finite state spaces. See however Friedman and Yellin (1996) and To (1995) for a generalization to continuous state spaces.

and some μ_i s.t. $(\mu_i, \mu_j) = \varphi^t(\omega)$. That is, $\delta_i^*(\omega)$ is the smallest endowment parameter that maximizes the fitness function $\tilde{F}_i(\cdot, \mu_j)$ along the flow originating in ω .

Lemma 3 *Suppose $\omega = (\mu_i, \mu_j)$ with $\mu_j([\hat{e}_j, \infty]) = m < 1$. Then there exists a $\delta_i^*(\omega) > 0$ such that for $\mathcal{I} := (0, \delta_i^*(\omega))$*

$$\frac{1}{\mu(\mathcal{I})} \int_{\mathcal{I}} \tilde{F}(e_i, \mu_j) d\mu_i(e_i) > \tilde{F}(0, \mu_j) > \tilde{F}(e', \mu_j), \forall e' \geq \hat{e}_i.$$

Proof To prove the first inequality note that due to growth monotonicity

$$\frac{1}{\mu(A)} \int_A \tilde{F}(e_i, \mu_j) d\mu_i(e_i) = 0 \leq \frac{1}{\mu(A')} \int_{A'} \tilde{F}(e_i, \mu_j) d\mu_i(e_i),$$

$\forall A \subset [\hat{e}_j, \infty]$, $A' \subset [0, \hat{e}_j)$. Therefore, $\mu_j^t([\hat{e}_j, \infty]) \leq m$, $\forall t \geq 0$.

Let \mathcal{M}_j be the set of all μ_j with the property that $\mu_j([\hat{e}_j, \infty]) \leq m$. We claim that \mathcal{M}_j is closed. Consider a sequence $(\mu_j)^n$ with $\lim_{n \rightarrow \infty} (\mu_j)^n = \mu_j^*$ and $(\mu_j)^n \in \mathcal{M}_j$ for all $n \in \mathbb{N}$. Under the sup-norm this implies that $\lim_{n \rightarrow \infty} |(\mu_j(A))^n - \mu_j^*(A)| = 0$ for all $A \in \mathcal{F}$. Since for all $n \in \mathbb{N}$ and $A \in \mathcal{F}$, $0 \leq (\mu_j(A))^n \leq 1$, $(\mu_j(\bigcup A))^n = 1$, and $(\mu_j([\hat{e}_j, \infty]))^n \leq m$, it follows that $\mu_j^* \in \mathcal{M}_j$. Hence, \mathcal{M}_j is closed.

By Corollary 1 it follows that $\delta_i(\mu_j) > 0$ for all $\mu_j \in \mathcal{M}_j$. Since \mathcal{M}_j is closed it follows that $\delta_i^*(\omega) > 0$ for all $\omega = (\mu_i, \mu_j)$ with $\mu_j \in \mathcal{M}_j$. By continuity of $\tilde{F}_i(\cdot, \mu_j)$, $\tilde{F}_i(e_i, \mu_j) > \tilde{F}_i(0, \mu_j)$ for all $0 < e_i < \delta_i^*(\omega)$. Hence, average fitness on \mathcal{I} must be strictly higher than $\tilde{F}_i(0, \mu_j)$.

The second inequality follows since by definition of \hat{e}_i , $\tilde{F}_i(e'_i, \mu_j) = 0$ for all $e'_i \geq \hat{e}_i$. However, $\tilde{F}_i(0, \mu_j) > 0$ for all μ_j with $\mu_j([0, \hat{e}_j]) > 0$. ■

Proof of Proposition 2 Fix some $\omega \in \bar{B}$. Based on ω we construct another state $\omega^0 := (\mu_1^0, \mu_2^0)$ by moving some probability mass from a mass point at 0 (if it exists) and some mass from $[\hat{e}_i, \infty]$ (if it exists) to the open interval \mathcal{I} defined in Lemma 3. To be precise let μ_i^0 , $i = 1, 2$ be such that

$$\mu_i^0(\{0\}) = \mu_i(\{0\}) - \varepsilon_1, \text{ with } \varepsilon_1 > 0 \text{ if } \mu_i(\{0\}) > 0$$

and

$$\mu_i^0([\hat{e}_i, \infty]) = \mu_i([\hat{e}_i, \infty]) - \varepsilon_2, \text{ with } \varepsilon_2 > 0 \text{ if } \mu_i([\hat{e}_i, \infty]) > 0.$$

The mass on \mathcal{I} is increased by setting

$$\mu_i^0(\mathcal{I}) = \mu_i(\mathcal{I}) + \varepsilon_1 + \varepsilon_2.$$

We can pick $\varepsilon_1 \geq 0$ and $\varepsilon_2 \geq 0$ such that $\mu_i^0(A) \geq 0, \forall A \in \mathcal{F}$ and such that $\mu_i^0(\cdot)$ becomes arbitrarily close to $\mu_i(\cdot)$ in the sup-norm. Payoff monotonicity in combination with Lemma 3 implies that

$$\hat{\mu}_i^0(\mathcal{I}) > \hat{\mu}_i^0(\{0\}) > \hat{\mu}_i^0([\hat{e}_i, \infty]), \quad (9)$$

where $\hat{\mu}_i^0(A) := \frac{\dot{\mu}_i^0(A)}{\mu_i^0(A)}$ is the growth rate of A .

Consider now some neighborhood U of ω with the property that, $\forall \omega' \in U, i = 1, 2$

$$\mu_i^0(\{0\}) + \mu_i^0([\hat{e}_i, \infty]) > \alpha \mu_i^0(\mathcal{I}), \alpha > 0. \quad (10)$$

We can choose ε_1 and ε_2 sufficiently small such that $\omega^0 \in U'$ for all open neighborhoods U' of ω contained in U . Lipschitz continuity together with (9) implies that any trajectory $\varphi^t(\omega^0)$ starting in ω^0 must eventually violate (10), i.e. it must leave U . ■

The importance of Proposition 2 lies in the fact that it rules out stable states in which some individuals have no endowment effect. On the other hand, stable states in which some individuals have very high endowment effect, which would make trade impossible, do not exist either. The non-existence of stable boundary states does not by itself exclude the possibility that the process spends a significant amount of time in the set of boundary states. However, the following proposition demonstrates that the proportion of individuals without or with very high endowment effects converges to zero as time progresses.

Let $D := \{\omega : \forall \delta > 0, \mu_i((0, \delta)) > 0, i = 1, 2\}$ denote the set of states in which in both population some probability mass is concentrated close to 0. Note that some $\omega \in D$ may be boundary states but they can never be stationary boundary states.

In the literature it is frequently assumed that initial states are interior in the sense that all types are represented in the initial population (see e.g. Cressman and Schlag, 1996, and Weibull, 1995). All such interior states certainly belong to D . In fact, for the next proposition it is sufficient that the states have full support on $(0, \hat{e}_i)$. Obviously, belonging to D is weaker still than this latter assumption.

Proposition 3 *Starting from any initial state $\omega^0 \in D$, the proportions of types in both populations with endowment parameters of 0 or larger than \hat{e}_i converge to 0 as $t \rightarrow \infty$.*

Proof By definition for all states $\omega^0 \in D$ there exists an interval \mathcal{I} as defined in Lemma 3. By Lipschitz continuity of the process it follows from (9) that for some $\kappa > 0$ and $A \in \{\{0\}, [\widehat{e}_i, \infty]\}$

$$\widehat{\mu}_i^t(A) - \widehat{\mu}_i^t(\mathcal{I}) < -\kappa \quad \forall t \geq 0, \quad i = 1, 2.$$

Thus,

$$\left(\frac{\widehat{\mu}_i^t(A)}{\widehat{\mu}_i^t(\mathcal{I})} \right) < -\kappa, \quad \forall t \geq 0$$

and hence

$$\frac{\mu_i^t(A)}{\mu_i^t(\mathcal{I})} < \frac{\mu_i^0(A)}{\mu_i^0(\mathcal{I})} e^{-\kappa t}, \quad \forall t \geq 0.$$

Since $\mu_i^0(\mathcal{I}) > 0$, the right hand side converges to zero as $t \rightarrow \infty$. Hence, $\mu_i^t(A) \rightarrow 0$ as well. ■

Proposition 3 is the main result of the paper. It demonstrates clearly that evolution favors individuals with an endowment effect. After a sufficiently long time almost all individuals will have a positive and finite endowment effect.

The question arises then whether the process ever settles down to a stable state. While we cannot prove the existence of a stable state for the general case, we can prove the existence of an evolutionary stable strategy combination (ESS) under the additional assumption that there always exists a unique $e_i^*(e_j)$ maximizing the reduced form fitness function $F_i(x^*(e_1, e_2), y^*(e_1, e_2))$. Selten (1980) has shown that in asymmetric contests an ESS must be a strict Nash equilibrium (and vice versa). A state $\omega = (\mu_1, \mu_2)$ is a strict Nash equilibrium of the population game if $\mu_i(\{e_i\}) = 1$, $i = 1, 2$ and

$$\widetilde{F}(e_i, \mu_j) > \widetilde{F}(e'_i, \mu_j), \quad \forall e'_i \neq e_i.$$

Proposition 4 *If the correspondences $H(e_1)$ and $H(e_2)$ are single valued, then there exists a strict Nash equilibrium ω^* , in which both populations are monomorphic.*

Proof By the Theorem of the Maximum (Berge, 1963) the correspondences $H(e_i)$ are upper hemi-continuous. If they are single valued, they are continuous reaction functions. We will argue that the reaction functions must cross. This follows from two facts. First, the reaction function $H(e_j)$ connects $H(e_j = 0)$ (which is strictly positive by Proposition 1) with $H(\widehat{e}_j) = 0$,

with $\widehat{e}_j < \infty$. Secondly, $H(\widehat{e}_j) > H(e_i = 0)$ by equation (3). Since this is true for both (continuous) reaction functions, $H(e_2)$ must cross $H(e_1)$ at least once from above. Let us call this point (e_1^*, e_2^*) . By definition of $H(e_i)$ it follows that (e_1^*, e_2^*) is a strict Nash equilibrium. ■

6 Conclusion

‘Learning to like what you have’ seems to be an evolutionarily successful strategy. We have shown in this paper that an apparent behavioral anomaly, the endowment effect, which has been observed in numerous experiments, can be explained by evolutionary arguments. We have argued that people acquire a preference for goods they own because it helps them in bargaining situations.

This is quite different from the observation that individuals may have strategic incentives to lie about their true preferences. As convincingly argued by Frank (1988) and Ellingsen (1996) it is not always possible to credibly signal preferences which one does not hold. In our setting individuals behave sincerely according to their preferences. Neither do they lie nor do they commit themselves to non-credible threats. They simply develop an endowment effect because individuals with an endowment effect end up with more resources and therefore higher fitness. Note, however, that overall the endowment effect causes an inefficiency since there is a suboptimal amount of trade. Feasible allocations which would be mutually beneficial in terms of fitness are not implemented due to the bias in preferences.

It is important to notice that once evolution has brought forth preferences with endowment effects individuals will reveal their endowment effects not only in bilateral trade but also in incentive compatible market situations. The systematic endowment effect observed in market experiments can neither be explained by strategic misrepresentation of preferences nor by erroneous behavior since in the latter case one should also expect to observe negative endowment effects, i.e. the case in which the average WTP is greater than the average WTA. Taking these considerations into account only explanations based on preferences with a ‘hard-wired’ endowment effects seem to be consistent with experimental data.

Several open questions remain. While we suppose that our results hold true for other cooperative bargaining solutions, this still has to be shown formally. It would also be of interest to consider more general formulations for the utility function. However, as pointed out above, including other

utility functions will make it only more likely that an endowment effect would result. Finally, we were not able to prove the existence of a stable state in which all individuals have some fixed endowment parameter for the general case. In our view this is nothing to worry about. We have shown that most of time most of the individuals will have some positive and finite endowment effect — possibly varying between individuals. Looking at behavior in the laboratory clearly reveals that one should not expect anything else. In nearly all experiments there is an enormous variety in subjects' behavior suggesting that theoretical results offering point predictions are doomed to fail.

References

- [1] BERGE, C. (1963), *Topological Spaces*, New York: MacMillan.
- [2] BOMZE, I. and PÖTSCHER, B. (1989), *Game Theoretic Foundations of Evolutionary Stability*, Berlin: Springer Verlag.
- [3] COURSEY, D., HOVIS, J., and SCHULZE, W. (1987), “The Disparity between Willingness to Accept and Willingness to Pay Measures of Value”, *Quarterly Journal of Economics*, 102, 679-690.
- [4] CARMICHAEL, L. and MACLEOD, B. (1996), “Territorial Bargaining”, mimeo, Queens University, Canada.
- [5] CHAMBERLAIN, E. (1948), “An Experimental Imperfect Market”, *Journal of Political Economy*, 56, 95-108.
- [6] CRESSMAN, R. and SCHLAG, K. (1996), “The Dynamic (In)Stability of Backward Induction”, SFB 303, Discussion Paper #B-347, Bonn University.
- [7] ELLINGSEN, T. (1996), “The Evolution of Bargaining Behavior”, mimeo, Stockholm School of Economics, forthcoming *Quarterly Journal of Economics*.
- [8] FRANK, R. (1988), *Passions with Reason*, New York: W.W. Norton.
- [9] FRIEDMAN, D. and YELLIN, J. (1996), “Evolving Landscapes for Population Games”, mimeo, University of California, Santa Cruz.

- [10] GALE, J., BINMORE, K., and SAMUELSON, L. (1995), "Learning to Be Imperfect: The Ultimatum Game", *Games and Economic Behavior*, 8, 56-90.
- [11] GÜTH, W. and YAARI, M. (1992), "An Evolutionary Approach to Explain Reciprocal Behavior in a Simple Strategic Game, in: *Explaining Process and Change — Approaches to Evolutionary Economics*, ed. U. Witt, Ann Arbor, 23-34.
- [12] HUCK, S. and OECHSSLER, J. (1996), "The Indirect Evolutionary Approach to Explaining Fair Allocations", mimeo, Humboldt University, Berlin.
- [13] KAHNEMAN, D., KNETSCH, J. and THALER, R. (1990), "Experimental Tests of the Endowment Effect", *Journal of Political Economy*, 98, 1325-1348.
- [14] KAHNEMAN, D., KNETSCH, J. and THALER, R. (1991), "The Endowment Effect, Loss Aversion, and Status Quo Bias", *Journal of Economic Perspectives*, 5, 193-206.
- [15] KANDORI, M., MAILATH, G. and ROB, R. (1993), "Learning, Mutation, and Long Run Equilibria in Games", *Econometrica*, 61, 29-56.
- [16] KNETSCH, J. (1989), "The Endowment Effect and Evidence of Non-reversible Indifference Curves", *American Economic Review*, 79, 1277-1284.
- [17] KNETSCH, J. and SINDEN, J. (1984), "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value", *Quarterly Journal of Economics*, 99, 507-521.
- [18] KNETSCH, J. and SINDEN, J. (1987), "The Persistence of Evaluation Disparities", *Quarterly Journal of Economics*, 102, 691-695.
- [19] KNEZ, M. and SMITH, V. (1987), "Hypothetical Valuations and Preference Reversals in the Context of Asset Trading", in Roth, A. (ed.): *Laboratory Experimentation in Economics: Six Points of View*, Cambridge: Cambridge University Press.
- [20] SELTEN, R. (1980), "A Note on Evolutionary Stable Strategies in Asymmetric Animal Conflicts", *Journal of Theoretical Biology*, 84, 93-101.

- [21] THALER, R. (1980), "Toward a Positive Theory of Consumer Choice", *Journal of Economic Behavior and Organization*, 1, 39-60.
- [22] TO, T. (1995), "Risk and Evolution", mimeo, University of St. Andrews.
- [23] VENTI, S. and WISE, D. (1990), "But They Don't Want to Reduce Housing Equity", in Wise, D. (ed.), *Issues in the Economics of Aging*. A NBER Project Report, Chicago: University of Chicago Press, 13-29.
- [24] WEIBULL, J. (1995), *Evolutionary Game Theory*, Cambridge, Mass.: MIT-Press.
- [25] WEINROBE, M. (1988), "An Insurance Plan to Guarantee Reverse Mortgages", *Journal of Risk and Insurance*, 55, 644-59.
- [26] YOUNG, P. (1993), "An Evolutionary Model of Bargaining", *Journal of Economic Theory*, 59, 145-168.