Reference-Dependent Preferences: Models and Experiments

By Maximilian Tenberge *

Abstract

In economic situations people form expectations prior to their decisions. These expectations represent a reference point and exert a strong influence on decision making and preferences. This paper surveys the theory of reference-dependent preferences. I will summarize the theoretical framework and present experimental evidence in support of preferences being reference-dependent. Additionally, I will address the still open and fundamental questions of how expectations are formed. That is: How are reference points formed?

Keywords: Expectation Formation, Reference-Dependent Preferences

*This paper is part of a seminar in Behavioral Economics at the University of Konstanz at the Chair of Applied Research in Economics of Prof. Dr. Urs Fischbacher. The author is a matriculated student at the University of Konstanz, Department of Economics, 78457 Konstanz: Maximilian Tenberge, E-mail:Maximilian.Tenberge@uni-konstanz.de.

1 Introduction

Reference-dependent models originate from prospect theory. In such models, utility is derived from comparing consumption levels to some reference levels. Yet, the definition of the reference level or the reference point was left imprecise or it was assumed that the status quo acts as the reference point as in Samuelson and Zeckhauser (1988). Against the view of Pesendorfer (2006) that reference points are free parameters to be determined by the researcher, Köszegi and Rabin (2006) argue that reference points are formed by recent expectations. In their concept, the reference point is endogenously determined as a function of the decision maker's probabilistic beliefs regarding the choice set he will face and his planned action for each possible choice set. Several empirical studies demonstrate that preferences are reference-dependent[1]. Frankly speaking, people do not evaluate an economic outcome based solely on how it ranks on an absolute scale, but also care about how it compares to relevant reference points.

Consider Johnny for instance. He is a very clever, foresighted and rational decision maker. Let him work for a Blue Chip company. He soon expects a raise of his salary from \$30K to \$50K. Unfortunately, the raise turns out to be not as high as expected. He only receives \$40K. Compared to his expectation this feels like a loss. However, Johnny also has to pay taxes. If he is lucky and only has to pay \$20K instead of \$30K in taxes, that certainly feels like a gain.

With expectations as a reference point, the framework of Köszegi and Rabin (2006) is applicable to many economic situations, as outlined in the following sections. In section 3 I will explain that Johnny's willingness to pay for a good is not fixed, but depends on the market environment and how he expects to respond to it. Particularly his valuation increases with the probability he believes to buy a good, and decreases if he had expected to buy it at lower prices. Section 4 will outline some experimental evidence in support of preferences being reference-dependent. Section 5 will focus on the still open question of how expectations and hence reference points are formed. Section 6 will conclude this paper.

2 A model of reference-dependent utility

The very first models with reference-dependence were introduced by Kahneman and Tversky (1979; 1991; 1992). They modeled decision-making with a value function that benchmarks gains and losses relative to a reference point. They argued that

[1]See e.g. Mas (2006), Heidhues and Köszegi (2008), Crawford and Meng (2008), Herweg et al. (2009).

losses weigh more heavily than same-sized gains and that the marginal contribution of gains and losses to the value function diminishes with the distance from the reference point. This builds the foundation of the reference-dependent model proposed by Köszegi and Rabin (2006). In their model, utility is not only derived from absolute consumption or wealth levels, but also from comparing the actual level to some reference level. The difference between the actual outcome and the reference level and the formation of the reference point are at crux of their model.

The model of Köszegi and Rabin (2006) formally starts from a given consumption level $c = (c_1, c_2, \ldots, c_K) \in \Re^K$ and a reference level of consumption $r = (r_1, r_2, \ldots, r_K) \in \Re^K$. They allow both the consumption level and the reference level to be stochastic. While the first is obvious, the latter satisfies that the reference level can be linked to beliefs about outcomes. The model allows the reference point to be a probability measure G over $\Re^K$. Recall our decision maker Johnny. Given such a probability measure G, Johnny calculates his overall utility of a consumption level c by comparing the consumption utility from c to each possible consumption level. Thus, his utility is given by

$$U(c|G) = \int u(c|r)dG(r).$$

If the consumption level c is as well drawn from a probability measure F over $\Re^K$, then Johnny's utility is given by

$$U(F|G) = \int \int u(c|r)dG(r)dF(c).$$

This takes into account that the sense of gains and losses Johnny will have from a given consumption outcome is derived by comparing it to every possible outcome under the reference lottery. Consider again the example from section 1. Johnny compares his received \$40K to the expected \$50K, which feels like a loss to him. He further compares his received \$40K to his prior salary of \$30K, which feels like a gain. Since loss aversion is a common feature among decision makers, Johnny's financial situation feels like an overall loss. However, his utility is not solely affected by the sensation of gains or the avoidance of losses, but also by the absolute level of his consumption. Köszegi and Rabin's (2006) approach explicitly takes into account that preferences depend on absolute levels, too. Their two-component utility function is hence defined over $\Re^K \times \Re^K \to \Re$ and takes the following form:

$$u(c|r) = \sum_{k=1}^{K} m_k(c_k) + \sum_{k=1}^{K} \mu(m_k(c_k) - m_k(r_k))$$

The first term $m_k(c_k)$ represents intrinsic consumption utility, where for each $k \in \{1, \ldots, K\}$, $m_k(\cdot)$ is a strictly increasing and differentiable function. It corresponds to the classical concept of physical outcome-based utility. The second term $\mu(m_k(c_k) - m_k(r_k))$ represents the reference-dependent gain-loss utility, where $\mu(\cdot)$ has the properties of the Kahneman and Tversky (1979) value function.

$A0$: $\mu(x)$ is continuous for all x, twice differentiable for $x \neq 0$, $\mu(0) = 0$.
$A1$: $\mu(x)$ is strictly increasing.
$A2$: If $y > x > 0$, then $\mu(y) - \mu(x) < -(\mu(-y) - \mu(-x))$
$A3$: $\mu''(x) \leq 0$ for $x > 0$ and $\mu''(x) \geq 0$ for $x < 0$
$A4$: $\frac{\mu'_-(0)}{\mu'_+(0)} \equiv \lambda > 1$, for $\mu'_+(0) \equiv lim_{x \to 0} \mu'(|x|)$ and $\mu'_-(0) \equiv lim_{x \to 0} \mu'(-|x|)$

$A0$ and $A1$ describe the general form of the function. $A2$ accounts for loss aversion for large stakes and $A4$ for small stakes. $A3$ captures the diminishing sensitivity of gains-loss utility. Additionally, $A3$ describes the kinked form of $\mu(\cdot)$. Below, figure 1 plots a representative form of the universal gain-loss function, with $x \equiv m(c) - m(r)$.

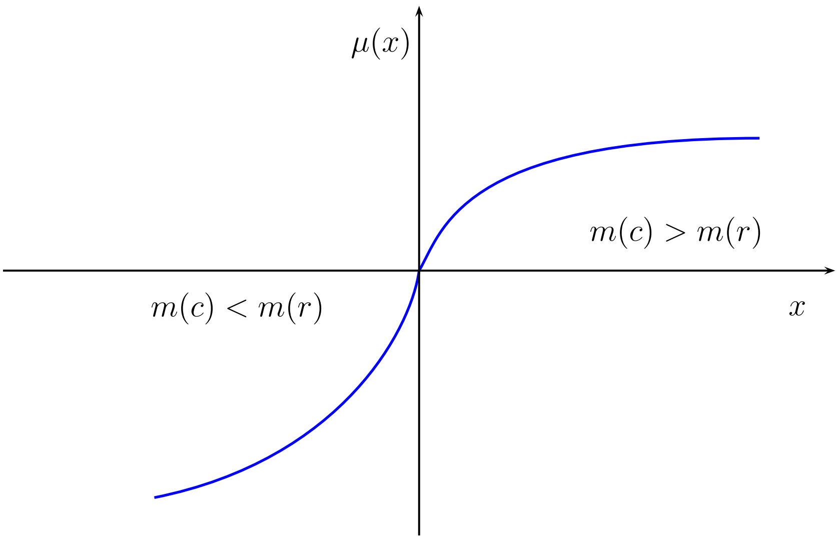

Figure 1. Universal gain-loss function

To solely analyze the implications of reference-dependence with loss aversion and to exclude effects of diminishing sensitivity, Köszegi and Rabin (2006) additionally assume a special linear case as sketched in figure 2. Their assumption $A3'$ states that for all $x \neq 0$, $\mu''(x) = 0$. Therefore, a gain-loss function satisfying $A3'$ takes the form of $\mu(x) = \eta x$ for gains and $\mu(x) = \eta \lambda x$ for losses. Here, $\eta \geq 0$ represents the weight Johnny attaches to gain-loss utility and $\lambda > 1$ represents Johnny's loss aversion coefficient. This linear functional form of the universal gain-loss function will be sufficient to explain several effects, as outlined in section 3. Köszegi and Rabin

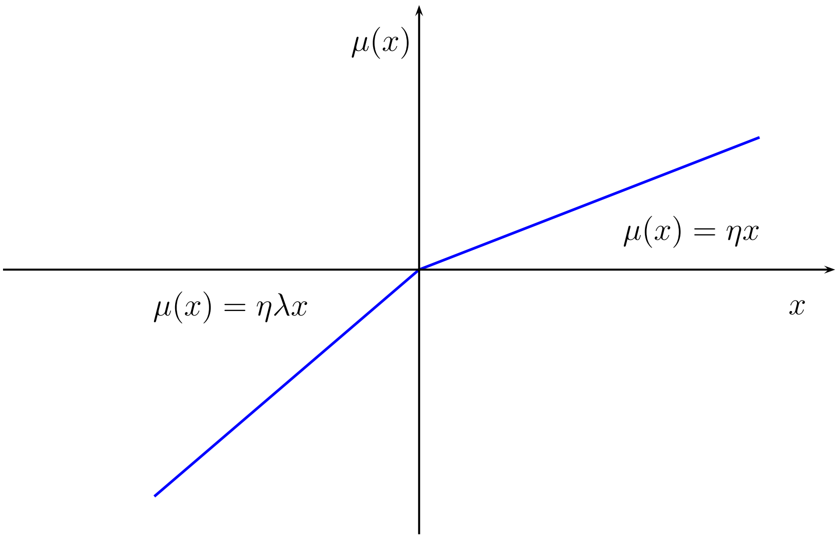

Figure 2. Special case: Linear gain-loss function

(2006) assume that consumption utility and gain-loss utility is separable over k dimensions. That is, Johnny evaluates consumption utility and gain-loss utility in each dimension separately by comparing c to each possible value that this dimension could take according to the reference lottery[2]. In combination with loss aversion, separability builds the foundations of important implications of reference-dependent utility, which will be discussed in section 3 and 4.

Based on Köszegi and Rabin's (2006) assumptions $A0 - A4$, the utility function replicates several properties[3]. Suppose the outcome c is fixed, then a lower reference point will make Johnny happier. This is obvious, if we remind us that utility in this model crucially depends on the difference between the actual outcome and the reference level. Thus, a smaller difference yields higher utility. A second property states that if Johnny forfeits his reference point for an alternative, he strictly prefers this alternative if that is his reference point. For small changes in utility, the function shares the qualitative properties of the standard formulation of prospect theory. This, however, is not true if the changes are large or occur quickly.

Beside formulating a specific utility function, Köszegi and Rabin (2006) also propose a model of reference point formation. Their model states that the reference point is determined by Johnny's beliefs about the choice sets he will face and the choice he will make from each set in the support of his beliefs. In such an environment, Johnny correctly predicts the environment and the corresponding choice sets he faces - e.g. a distribution of prices - and his own reaction to each choice set,

[2]Further details on separability are discussed in Köszegi and Rabin (2007).
[3]For a formal definition of the propositions stated here and their proofs, the reader may refer to the original paper of Köszegi and Rabin (2006).

e.g. a buying decision. Taking the reference point generated by these expectations as given, Johnny then plans for each contingency by maximizing expected utility. Formally, his beliefs are described by a distribution Q over $\Re$ which captures the possible choice sets $\{D_l\}_{l\in\Re}$, with $D_l \subset \Delta(\Re^K)$. Köszegi and Rabin (2006) define then a *personal equilibrium* as a selection $\{F_l \in D_l\}_{l\in\Re}$ if for all $l \in \Re$ and $F_l^{'} \in D_l$, $U(F_l | \int F_l dQ(l)) \geq U(F_l^{'} | \int F_l dQ(l))$ is fulfilled. If he expects the distribution of outcomes $\int F_l dQ(l)$ and expects to choose F_l from D_l, then the above definition states that Johnny will indeed choose F_l from the choice sets D_l. This concept of personal equilibrium is most related to the notion of loss aversion equilibrium in Shalev (2000). That is a Nash equilibrium defined for multiplayer games in which each player's reference point is fixed, such that the reference point equals the players implicitly defined reference-dependent expected utility.

In Köszegi and Rabin's (2006) framework, it is possible that Johnny has multiple personal equilibria. But since he has a single utility function, he can rank the outcome of each possible personal equilibria in terms of ex-ante expected utility. Principally, Johnny is free to make any plan as long as he follows it through. But finally, he will choose the most preferred plan. In this case, Johnny chooses a *preferred personal equilibrium*, which is a selection $\{F_l \in D_l\}_{l\in\Re}$ and a personal equilibrium and fulfills $U(\int F_l dQ(l) | \int F_l dQ(l)) \geq U(\int F_l^{'} dQ(l) | \int F_l^{'} dQ(l))$ for all personal equilibrium selections $\{F_l \in D_l\}_{l\in\Re}$.

To illustrate the above mentioned equilibrium concepts, consider Johnny at a time of a purchase decision. At $t = 0$, Johnny starts to focus on the decision and knows all the possible choice sets he might face including the probabilities with which they might occur. At $t = 1$, Johnny plans for each possible choice set. In our case, whether or not he will buy at the given prices. He thereby determines his reference lottery. At $t = 2$, the choice set Johnny finally will face is determined by nature. He follows through with his plan for this choice set which then results in a deterministic or probabilistic outcome. The personal equilibrium concept requires Johnny's action in $t = 2$ to be optimal given his beliefs from $t = 1$, his beliefs in $t = 1$ to be rational and his plans to be credible, such that he has no incentive to deviate from them. This can be viewed as similar to subgame perfection in multi-player game theory. The personal equilibrium concept requires Johnny's planned behavior to be optimal given his reference point, and the preferred personal equilibrium concept requires Johnny to choose the personal equilibrium with the highest ex-ante expected utility.

Having outlined the theoretical foundations of the model, I will emphasize some implications of the model in the following section.

3 Implications of the model

In their original paper, Köszegi and Rabin (2006) apply their model to labor supply and consumption decisions. In the context of labor supply, their model's predictions are consistent with the income targeting hypothesis[4]. In the context of consumption decisions their model makes new predictions. In addition to the endowment effect, Köszegi and Rabin (2006) predict an attachment effect and a comparison effect.

To outline the implications of Köszegi and Rabin's (2006) model, I will first describe a formal framework by continuing to analyze our decision maker, Johnny. Suppose Johnny derives utility from a consumption bundle $(c_1, c_2) \in \Re^2$, with c_1 representing ownership of a widget and c_2 representing Johnny's money wealth. Consider the initial endowment of Johnny as $(0, 1)$, his utility function $m(c_i)$ as being linear in wealth, and the gain-loss function as described by assumption $A3'$. Consider further that Johnny faces the problem to determine the maximum price he would be willing to pay for a widget, and the minimum price for which he would be willing to sell it. Assume that in this environment his reference point Ω is a lottery in which he receives the widget with probability q and does not receive it with probability $1-q$ [5]. A situation that gives justification to such a reference point could occur in an economic experiment. In such, Johnny - as a participant - could expect with some probability that he will be endowed by the experimenter with a widget. If he is endowed with it, he could be given the chance to sell it. If he is not endowed with it, he could buy it. When faced with the purchase situation, Johnny will compare his utility from buying at a price p to his utility from not buying at this price and the reference point Ω. The maximum price he is willing to pay (WTP) is then determined by the price that makes him indifferent between buying and not buying. Hence, Johnny compares:

$$U(\textit{not buy at } p \mid \Omega) = U(\textit{buy at } p \mid \Omega)$$

$$m_1(0)+w+q\eta\lambda(m_1(0)-m_1(1)) = m_1(1)+(w-p)+(1-q)\eta(m_1(1)-m_1(0))+\eta\lambda(-p)$$

[4] See Camerer et al. (1997), Goette et al. (2004), and Fehr and Goette (2007). Under the income targeting hypothesis, workers set a short-term income target that serves as a reference point. The hypothesis predicts that workers supply more labor when wages are low and less when high, in contrast to classical predictions that workers should supply more labor when wages are high. Köszegi and Rabin's (2006) model makes classical predictions when wages are as expected, but follows the income targeting hypothesis hypothesis if wages are a surprise.

[5] This reference point can be considered as exogenously given - e.g. during an experiment - or as a personal equilibrium in which Johnny expects to face the lottery Ω with probability $(1-\alpha)$ and some alternative choice D with probability α. If α approaches 0, the contribution of outcomes from D to gain-loss utility will be minimal. Hence, for very small α Johnny will evaluate choices from D as if Ω is the reference point.

$$WTP = p = \frac{1+\eta+q\eta(\lambda-1)}{1+\eta\lambda}(m_1(1)-m_1(0))$$

In the opposite case, the minimum price at which Johnny is willing to sell (WTA), is determined by the price that makes him indifferent between selling and not selling. Here, he compares:

$$U(not\ sell\ at\ p \mid \Omega) = U(sell\ at\ p \mid \Omega)$$

$$m_1(1)+w+(1-q)\eta(m_1(1)-m_1(0)) = m_1(0)+(w+p)-q)\eta\lambda(m_1(1)-m_1(0))+\eta(p)$$

$$WTA = p = \frac{1+\eta+q\eta(\lambda-1)}{1+\eta}(m_1(1)-m_1(0))$$

Figure 3 plots the WTP in blue and the WTA in red, for $\lambda = 2$ and $\eta = 0.5$, while assuming that $m_1(1) - m_1(0) = 1$.

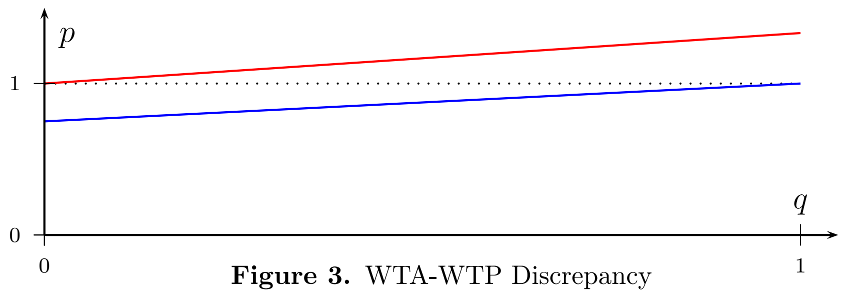

Figure 3. WTA-WTP Discrepancy

Lets first look at the slope of both functions. Both are increasing with q. For WTP the intuition is as follows: The higher the probability of receiving the widget the stronger Johnny feels attached to it. Receiving something with probability 0.99 feels close to already possessing it. Not receiving it consequently results in a feeling of loss. To prevent that possible loss, Johnny may be willing to pay a higher price. The intuition for WTA is similar: The higher the probability of selling the widget the stronger Johnny feels attached to the money he may receive. This attachment effect has interesting implications for business situations. Imagine Johnny wants to buy a new car. The car salesperson may offer a relatively low price for a car that Johnny then accepts. After letting Johnny feel comfortable with his final purchasing decision, the salesman comes back from a talk with the manager. Surprise! The low price was not accepted by the management and a higher price is proposed. By using the tactic of throwing a lowball, the car salesperson aims to create a strong attachment effect in order to increase Johnny's willingness to pay.

Going back to the functional form of the valuation, we can find a second implication. If we fix the probability of getting a widget, and the price Johnny expects to pay decreases - e.g. he expects to bargain a discount - paying a higher price then feels more like a loss. This comparison of prices Johnny expects to prices that induce purchase decreases his willingness to pay higher prices.

If we now look at the difference between WTA and WTP, we see that it only depends on the denominator and more specifically on the loss aversion coefficient λ. The difference between these two valuations of the same widget stems from the fact, that in one situation (WTA) Johnny owns the widget and in the other situation (WTP) Johnny does not own the widget. If an owner is asked to sell his widget but expects to be endowed with it in the future he has to be compensated not only for the intrinsic value of the widget but also for the fact that he has to forgo his reference point, which he certainly does not prefer.

4 Experimental applications

Since Köszegi and Rabin's (2006) seminal paper, several studies complemented the research on reference dependent preferences, particularly falsifying how expectations determine reference points[6]. Prior to Köszegi and Rabin (2006), the reference point has often been taken to be the status quo, as in Samuelson and Zeckhauser (1988). That has not changed with reference points determined by expectations. Modern theories now assume that individuals among other things might expect to keep their current status. That argument is consistent with the status quo as a reference point and recent expectations as a reference point.

Several studies investigated the status quo as a reference point. Among them, experiments by Knetsch (1989) and Plott and Zeiler (2007). Knetsch (1989) found that only 10 percent of participants in an experiment who where endowed with an item were willing to exchange it for another item. Standard theory would predict that about half the participants should exchange. This exchange asymmetry was interpreted as resulting from loss aversion around the status quo. However Knetsch's (1989) experiment did not account for the expectations participants in his experiment may have had. Using the framework of Köszegi and Rabin (2006) would suggest that participants expected to leave the experiment with the endowed item, until the opportunity of exchange is explicitly offered. In this case endowment and expectations coincided.

Plott and Zeiler (2007) argued that Knetsch's (1989) results were due to the experimental procedure he used. By altering the experimental setting of Knetsch (1989) they showed that the endowment effect can disappear if certain experimental

[6]Studies with expectations based reference point include Mas (2006), Heidhues and Köszegi (2008), Crawford and Meng (2008), Herweg et al. (2009).

procedures are used. They interpreted their results in favor of classical preferences interacting with experimental settings. Nevertheless, their results were also consistent with the endowment effect driven by reference-dependent preferences. If they had accounted for the expectations participants could have, they may have concluded that the reference-point is determined by expectations and not by endowment.

Taking into account the importance of expectations for determining the reference point, Ericson and Fuster (2009) applied an experimental design in which they exogenously induced expectations[7]. They endowed participants with a mug and randomized the probability that the participant will be able to exchange the mug for an alternative (a pen). They grouped their participants into a low treatment - being able to exchange with low probability $p = 0.1$ - and a high treatment - being able to exchange with high probability $p = 0.9$. After being seated in the laboratory and having read the instructions, participants answered a filler questionnaire and then had to indicate if they would like to exchange the mug for the pen, if they were given the chance to. This chance depended on the exogenously assigned probability.

Before commenting on the experimental results of Ericson and Fuster (2009), I will outline the formal mechanism in their experiment, which will give additional insights on how the model of Köszegi and Rabin (2006) works. Recall our decision maker Johnny who participates in this experiment. In the beginning he is endowed with a mug that yields a utility of u_{mug}. If he plans to exchange the mug for a pen, he can expect with probability p to be allowed to exchange and will then be endowed with a pen, which then is his reference point $r_{end\ w/pen}$. Conversely, with probability $1 - p$ he will not be allowed to exchange, and hence will keep the mug. In that case his reference point is $r_{end\ w/mug}$. Given this expectations, the utility of Johnny of an outcome c is then

$$U(c|Exchange) = pu(c|r_{end\ w/pen}) + (1-p)u(c|r_{end\ w/mug}).$$

The expected utility of Johnny, if he sticks to his plan to exchange is then given by:

$$\begin{aligned} EU(Exchange|Exchange) = p[u_{pen} + (1-p)\eta(u_{pen} - \lambda u_{mug})] \\ + (1-p)[u_{mug} + p\eta(u_{mug} - \lambda u_{pen})] \end{aligned}$$

If Johnny planned to exchange, but then deviates from his planned decision by

[7] If expectations are induced during experiments, researcher have to avoid that subjects make wrong inferences - e.g. from exogenous probabilities. Ericson and Fuster (2009) transparently randomized treatments to make clear to the participants that no information concerning the value of an item is included in the probability.

keeping the mug, his expected utility from doing so, would be:

$$EU(Keep|Exchange) = u_{mug} + p\eta(u_{mug} - \lambda u_{pen})$$

Hence, choosing to exchange is then a personal equilibrium if:

$$EU(Exchange|Exchange) \geq EU(Keep|Exchange)$$

That leads to the following condition:

$$u_{pen} \geq u_{mug} \underbrace{\frac{1 + \eta(\lambda + p(1-\lambda))}{1 + \eta(1 - p(1-\lambda))}}_{x(p)}$$

The functional form of $x(p)$ is as follows: $x'(p) < 0$, $x(0) = \frac{1+\eta\lambda}{1+\eta} > 1$, $x(\frac{1}{2}) = 1$, $x(1) = \frac{1+\eta}{1+\eta\lambda} < 1$. If we now assume a parametrization as in section 3, with $\lambda = 2$ and $\eta = 0.5$, we can illustrate the behavior of $x(p)$ in figure 4. This means, if

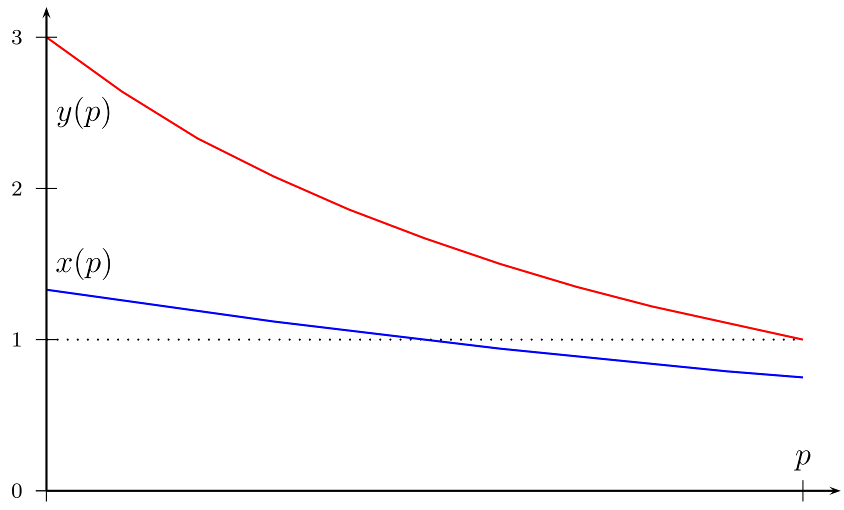

Figure 4. Functional form of x(p) and y(p)

the probability of having the chance to exchange the mug for the pen is low - in this case $p < \frac{1}{2}$ - then u_{pen} must be much higher for exchange to be a personal equilibrium. Interestingly, for $p > \frac{1}{2}$ to exchange the mug for the pen might be a personal equilibrium even if u_{pen} is lower than u_{mug}. For exchange to be optimal behavior, it has also to fulfill the preferred personal equilibrium condition. That is, $EU(Exchange|Exchange) \geq EU(Keep|Keep)$, which result in the following condition.

$$u_{pen} > u_{mug} \underbrace{\frac{1 - \eta(1-p)(1-\lambda)}{1 + \eta(1-p)(1-\lambda)}}_{y(p)}$$

As we can see from figure 4, to exchange the mug for the pen can only be a preferred personal equilibrium if u_{pen} is larger than the consumption utility of a mug. As p increases, to exchange becomes a preferred personal equilibrium for more pairs of u_{pen} and u_{pen}, because the utility surplus from a pen becomes smaller.

The results of Ericson and Fuster (2009) are as follows. They found that the probability a participant will be allowed to exchange significantly influences his willingness to do so. Participants in the high treatment were one and a half times more likely to choose to exchange than participants in the low probability treatment. They further found that more participants in the low treatment liked the mug significantly more than in the high treatment. Both results can not be explained by standard theory, which would predict that people in both treatment are equally likely to like the mug more than the pen. In addition and according to standard theory, the decision to exchange should not depend on the probability an individual will be allowed to do so, but on the sole argument whether the exchange is advantageous or not.

This experiment nicely confirms that reference points matter for choice and that reference points are determined by expectations.

5 Reference point formation

In the previous sections, I outlined the theoretical foundations and implications of reference-depended preferences based on the model of Köszegi and Rabin (2006). Additionally, I have presented experimental evidence that support the reference-dependence of preferences. Although the idea of reference-dependence has well settled into behavioral economics, the exact definition of a reference point and its development over time is still an open question. While in some studies, particular in a marketing context, the reference point is interpreted as the consumer's expectation of prices or quality, other studies refer to the reference point as a weighted average of past consumption or a habitual level of consumption[8]. Yet, other studies regard the reference point as a mixture of an aspiration level and some critical survival point or as the weighted average of past stimuli[9]. All these studies have in common that they rarely explain how the reference point is formed or updated. An important questions is how the reference point is updated as a function of past decisions and

[8]For studies picking up these two notions of the reference point see Lattin and Bucklin (1989), Putler (1992), Hardie et al. (1993), Bell and Bucklin (1999), Bell and Lattin (2000), Popescu and Wu (2007), Wathieu (2004).

[9]See March (1988), March and Shapira (1992), Helson (1964), Lim (1995), Kopalle et al. (1996), Chen and Rao (2002), Compte and Jehiel (2003).

outcomes. If we go back to the initial statement that models of reference-dependence originate from prospect theory, we are reminded that prospect theory in addition to the value function, uses a probability weighting function.

Baucells et al. (2008) borrow the idea of a probability weighting function to propose a model of reference point formation[10]. Their premise is that the reference point is the outcome of a mental process in which all available information are used and weighted. In their model the reference point is a function of available information, which are weighted using a information weighting function.

In the following, I will again rely on the help of Johnny to outline their model. Consider a rank dependent model as in cumulative prospect theory Kahneman and Tversky (1992) in which prospects are evaluated using two steps. Baucells et al. (2008) replace the vector of prospects by a price sequence $(y_1, y_2, \ldots, y_n)$, which essentially is a vector of information. In contrast to decisions under uncertainty, this vector is sorted along the timeline and not from small to big. In the first step, Johnny sorts the information from most recent past to most distant past, such that $(y_1^*, y_2^*, \ldots, y_n^*) = (y_n^*, y_{n-1}^*, \ldots, y_1^*,)$[11]. Let $(p_1^*, p_2^*, \ldots, p_n^*)$ be the corresponding probabilities. Then, in the second step, Johnny accumulates the probabilities, so that $(p_1^*, p_1^* + p_2^*, \ldots, 1) = (P_1^*, P_2^*, \ldots, P_n^*)$. Baucells et al. (2008) assume the simplest weighting, resulting in: $(P_1^*, P_2^*, \ldots, P_n^*) = (\frac{1}{n}, \frac{2}{n}, \ldots, 1)$. In the third step the information weighting function is introduced which transforms $(P_1^*, P_2^*, \ldots, P_n^*)$ into $(w(\frac{1}{n}), w(\frac{2}{n}), \ldots, w(1))$. Next, the information weights are calculated in the following way: $\pi_i^* = w(\frac{n-i+1}{n}) - w(\frac{n-1}{n}), i = 1, \ldots, n$. If the linear case of the value function, as in section 2, is assumed, the reference point in period m is given by[12]:

$$RP_m = \sum_{i=1}^{m} \pi_i y_i \ ; \ 1 \leq m \leq n$$

This function states that Johnny's reference point is a weighted average of past information. He forms his reference point by sorting information from recent to past, and weights these information using an inverse s-shaped function. Baucells et al. (2008) stick to Köszegi and Rabin's (2006) argument that the reference point consist of recent beliefs. Soon after Johnny received the last piece of information y_m, he will form his reference point. In Baucells et al. (2008) finance context, a new piece of information y_{m+1} - e.g. a new asset price - will be perceived as a gain if it is

[10]Even though, their approach is explicitly modeled in a finance context, it can be used to explain the general properties of expectations formation.

[11]Here, Baucells et al. (2008) use $y_i^* = y_{n-1+1}$.

[12]If a non-linear value function is assumed, the reference point would be the certainty equivalent of the weighted average value of each price.

higher than RP_m and as a loss if otherwise. If w is assumed to be inverse s-shaped, the first and the last piece of information influence the reference point the most. In this model, the first piece of information will receive the weight $1 - w(\frac{n-1}{n})$ and the last will receive $w(\frac{n-1}{n})$. The remaining weight $w(\frac{n-1}{n}) - w(\frac{1}{n})$ is then distributed among the $n - 2$ remaining pieces of information.

Baucells et al. (2008) experimentally tested their theory of reference point formation[13]. They estimated the reference point by asking subjects which selling price would make them neither happy nor unhappy after they observed a stock price sequence. By regressing the reference point on the purchase price, the current price, the intermediate prices, the highest and lowest price, and the average and recency-weighted average price, Baucells et al. (2008) showed that the reference point is most influenced by the first and the last observed stock price. In their information sequences of length 3 to 10, the purchase price received a weight of 50%, the current price a weight of 30% and the average of intermediate prices a remaining weight of 20%.

Beside Baucells et al. (2008), several other researcher investigated the formation of reference points. However, only Baucells et al. (2008) proposed a theoretical mechanism to explain how the reference point is actually formed and updated. Chen and Rao (2002), for example, found that the order in which two equally sized, but differently signed events - e.g. same-sized gains and losses - influence the subject's final affective state. That suggests that a shift in the reference point must have occurred after the first event. Additionally, they found that adding a time lapse between the two events generated results consistent with a greater shift in reference points. Gneezy (2005) endowed participants in his experiments with a stock and asked them about their willingness to hold or sell the stock as its price varied over several trading periods. He assumed that subjects willingness to sell are the highest when the current price is equal to the reference point. He showed that assuming the peak price to be the reference point best explained subjects' willingness to sell that stock.

All of these studies suggest that the reference point is path-dependent. Past information - e.g. stock prices - additional to more recent information have significant impact on the current reference point. Yet, these studies lack some important criteria. Chen and Rao's (2002) study does not allow them to estimate the location of a new reference point, and Gneezy's (2005) method can only estimate the reference point relative to prior stock prices if the subject sold the stock. While Baucells et al.

[13]For a detailed outline of their experimental design, the reader may refer to Baucells et al. (2008) section 3.

(2008) propose a formal method of reference point formation, none of these studies estimate the exact magnitude of reference point adaptation. Future research may determine how reference points exactly move after gains and losses.

6 Discussion

While the theory of reference-dependent preferences has established itself in economic research, the fundamental question of how expectations are formed remains open. The timing and the state of the reference lottery play an important role. Köszegi and Rabin's (2006) model uses the premise that preferences depend on expectations formed after the decision maker started focusing on a decision. However, the timing effect of expectation forming is crucial. If Johnny had been thinking about a purchase for a long time, his expectations from before he knew the price will affect his preferences. The appropriate reference lottery is then the lottery representing his probabilistic beliefs over prices. But if he only considered the possible purchase upon seeing the good at some store, his reference lottery should be the deterministic lottery corresponding to the relevant price at that time. Here Köszegi and Rabin's (2006) model lacks a clear guide to the timing of expectation formation. A further important question is the way consumption outcomes are compared to reference levels. While in Köszegi and Rabin (2006) each outcome is compared to all outcomes in the support of the reference level, in Sugden (2003) an outcome is compared only to the outcome that would have resulted from the reference lottery in the same state. This is an important difference, if we consider that the reference point may include some information about forgone decisions. Particular, in a context of investment decisions, outcomes might only be compared to missed decisions rather than to all possible ones. This may link regret theory to a model of reference-dependent preferences.

References

BAUCELLS, M., M. WEBER, AND F. WELFENS (2008): "Reference Point Formation and Updating: Theory and Experiments," *Working Paper, University of Mannheim.*

BELL, D. R. AND R. E. BUCKLIN (1999): "The role of internal reference points in the category purchase decision," *Journal of Consumer Research*, 26, 128–143.

BELL, D. R. AND J. M. LATTIN (2000): "Looking for loss aversion in scanner panel data: The confounding effect of price response heterogeneity," *Marketing Science*, 19, 185–200.

CAMERER, C., L. BABCOCK, G. LOEWENSTEIN, AND R. THALER (1997): "Labor supply of New York City cabdrivers: One day at a time," *The Quarterly Journal of Economics*, 112, 407–441.

CHEN, H. AND A. R. RAO (2002): "Close Encounters of Two Kinds: False Alarms and Dashed Hopes," *Marketing Science*, 21, 178–196.

COMPTE, O. AND P. JEHIEL (2003): "Bargaining with reference dependent preferences," *Working Paper, Paris School of Economics.*

CRAWFORD, V. P. AND J. MENG (2008): "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependence Preferences with Rational-Expectations Targets for Hours and Income," *Working Paper, University of California, San Diego.*

ERICSON, K. M. M. AND A. FUSTER (2009): "Expectations as Endowments: Reference-Dependent Preferences and Exchange Behavior," *Working Paper, Harvard University.*

FEHR, E. AND L. GOETTE (2007): "Do workers work more if wages are high? Evidence from a randomized field experiment," *American Economic Review*, 97, 298–317.

GNEEZY, U. (2005): "Updating the reference level: Experimental evidence," *Experimental business research*, 263–284.

GOETTE, L., D. HUFFMAN, AND E. FEHR (2004): "Loss aversion and labor supply," *Journal of the European Economic Association*, 2, 216–228.

Hardie, B. G. S., E. J. Johnson, and P. S. Fader (1993): "Modeling loss aversion and reference dependence effects on brand choice," *Marketing Science*, 12, 378–394.

Heidhues, P. and B. Köszegi (2008): "Competition and price variation when consumers are loss averse," *American Economic Review*, 98, 1245–1268.

Helson, H. (1964): *Adaptation-level Theory*, Harper & Row New York.

Herweg, F., D. Müller, and P. Weinschenk (2009): "Binary Payment Schemes: Moral Hazard and Loss Aversion," *American Economic Review, forthcoming.*

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–292.

——— (1991): "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *The Quarterly Journal of Economics*, 106, 1039–1061.

——— (1992): "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Knetsch, J. L. (1989): "The endowment effect and evidence of nonreversible indifference curves," *The American economic review*, 79, 1277–1284.

Kopalle, P. K., A. G. Rao, and J. L. Assuncao (1996): "Asymmetric reference price effects and dynamic pricing policies," *Marketing Science*, 15, 60–85.

Köszegi, B. and M. Rabin (2006): "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121, 1133–1165.

——— (2007): "Reference-Dependent Risk Attitudes," *American Economic Review*, 97, 1047–1073.

Lattin, J. M. and R. E. Bucklin (1989): "Reference effects of price and promotion on brand choice behavior," *Journal of Marketing Research*, 26, 299–310.

Lim, R. G. (1995): "A range-frequency explanation of shifting reference points in risky decision making," *Organizational Behavior and Human Decision Processes*, 63, 6–20.

March, J. G. (1988): "Variable risk preferences and adaptive aspirations," *Journal of Economic Behavior & Organization*, 9, 5–24.

March, J. G. and Z. Shapira (1992): “Variable risk preferences and the focus of attention,” *Psychological Review*, 99, 172–183.

Mas, A. (2006): “Pay, reference points, and police performance,” *NBER Working Paper*.

Pesendorfer, W. (2006): “Behavioral Economics comes of age: A Review Essay on ”Advances in Behavioral Economics”,” *Journal of Economic Literature*, 44, 712–721.

Plott, C. R. and K. Zeiler (2007): “Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory?” *American Economic Review*, 97, 1449–1466.

Popescu, I. and Y. Wu (2007): “Dynamic pricing strategies with reference effects,” *Operations Research*, 55, 413.

Putler, D. S. (1992): “Incorporating reference price effects into a theory of consumer choice,” *Marketing Science*, 11, 287–309.

Samuelson, W. and R. Zeckhauser (1988): “Status quo bias in decision making,” *Journal of Risk and Uncertainty*, 1, 7–59.

Shalev, J. (2000): “Loss aversion equilibrium,” *International Journal of Game Theory*, 29, 269–287.

Sugden, R. (2003): “Reference-dependent subjective expected utility,” *Journal of Economic theory*, 111, 172–191.

Wathieu, L. (2004): “Consumer Habituation,” *Management Science*, 50, 587–596.