Reference price distortion*

József Sákovics

University of Edinburgh

December 5, 2007

Abstract

I show that when consumers (mis)perceive prices relative to reference prices, budgets turn out to be soft, prices tend to be lower and the average quality of goods sold decreases. These observations provide explanations for decentralized purchase decisions, for people being happy with a purchase even when they have paid their "valuation", and for why trade might be detrimental to welfare.

1 Introduction

This paper is intended as a building block in the research programme started by Richard Thaler (1980) towards a behaviorally based theory of consumer choice. As I am wary of imposing ad hoc behavioral assumptions, I use the mere existence of a single, empirically well-established[1] trait of consumer behavior to provide an explanation for a series of stylized facts and to make some further testable predictions.

*I thank Randy Bucklin, Botond Kőszegi, Carmen Matutes and Peter Sinclair, as well as audiences at the ESEM 2007, the 2007 mini-conference on "Behavioural Economics" in Edinburgh and the Theory seminar in Manchester for useful comments.

[1]See Mazumdar et al. (2005) for a review of the literature.

The observation underlying my analysis is that consumers have a prior idea of what a certain item "would" cost, and this *reference price* affects their purchasing decision. I do not want to theorize much about where this value comes from, though some kind of "rational" expectation is a clear contender. The conceptual novelty of my approach is that instead of assuming that the reference price affects *preferences*,[2] I posit that consumers behave as if they had standard preferences but had a distorted view of observed *prices*. One could view this as a straightforward error in price perception, but I prefer to think of it as a proxy for a more complex psychological phenomenon, which nonetheless leads to the same behavior and has the same welfare consequences. The nature of the misperception is simple: any price that is above the reference price seems even higher, and any price below the reference price seems even lower than it actually is. It is as if consumers were wearing glasses, where the concavity/convexity of the lenses was determined by whether the reference price is higher or lower than the price currently observed. As a further distinction from the literature, I do not assume that consumers are loss averse (though I allow for it), not even in the limited sense of price distortions being higher upwards than downwards.

Even before turning to market applications, one can use the above set-up to say something revealing about people's willingness to pay for an object. Depending on the relative position of the price to the reference price, one can be willing to spend more than one's intrinsic valuation of the object, or alternatively, strictly less than that. This observation is powerful in itself but, more importantly it can lead to a host of further implications when our agents are involved in a more complex interaction. Some of these I explore in this paper others are left for future research.

In the context of a market for a single homogeneous good, I show that equilibrium

[2] This has been the approach taken by all the theoretical studies, e.g. Thaler (1985), Putler (1992), and more recently, Ariely et al. (2004), Heidhues and Kőszegi (2005, 2006) and Kőszegi and Rabin (2006). Winer (1986) sidesteps this issue by directly modelling the dependence of the probability of purchase on the difference between the actual and the refernce price. While useful for regression analysis, this approach does not lead to a full-fledged theory of consumer behavior.

prices in Cournot markets would tend to be lower than in the neoclassical model, for any market structure except perfect competition. This result also generalizes to horizontally differentiated goods, since it is driven by the increase in the price elasticity of demand as a result of the price distortion.

When there is vertical differentiation, I show that the loss of market share of a high quality incumbent to a low quality entrant is increased as a result of reference price dependence. This finding provides a (partial) explanation of why local high quality firms may be driven out of the market by cheap low quality imports even when this is not efficient.

Next, I turn to the optimal purchase quantity problem. Here I show that – if some reasonable sufficient condition holds – the amount spent will vary in a somewhat counter-intuitive way: when the price is below (above) the reference price the consumer spends more (less) than if she correctly predicts the price. Note that this is without assuming that stockpiling is possible. When I incorporate this finding into the standard consumer choice problem of choosing the optimal consumption bundle given a budget constraint, it leads to notional/soft budgets,[3] as they are satisfied not for the true prices, rather the adjusted ones. I argue that this is not unreasonable, especially in a situation of sequentially distributed choice.

Finally I review some of the relevant empirical literature and conclude.

2 The reference-adjusted price

Based on the evidence cited above, my first axiom posits the existence of a reference price.

[3]Deaton (1977) has nearly hit on the idea of soft budget constraints in a model where there is inflation but the consumers only observe the price increase of a single good at a time. If they (wrongly) believe that the prices of the other goods have stayed constant, they will spend less than optimal (and less than mentally budgeted) on the good at question. He then goes on to look at the aggregate behavior of the economy and does not follow up the consequences on individual behavior.

Axiom 1 *Consumers have a reference price attached to each commodity. Different consumers – or the same consumer at different times – need not attach the same reference price to the same commodity.*

Taking the existence of reference prices as given, the question is how do they influence the consumers' purchasing decisions. One road to go down could be to assume that the reference price (together with the actual price) enters the consumer's utility function. This is the road existing theoretical approaches have taken (c.f. footnote 2). I make the alternative assumption that neither the actual nor the reference price affect the utility derived from purchasing an item.[4] Instead, all the action takes place via an adjustment of the actual price, which then modifies consumer behavior exclusively through its effect on expenditure, without affecting preferences (assuming quasi-linear utility). The literal interpretation for the price adjustment is that the consumer *misperceives* the actual price. However, I am not making a physiological/psychological statement here (as that is not my expertise). Rather, my claim is that consumers behave *as if* they were unable to accurately observe prices. As this approach does not interfere with preferences, it imposes strict discipline on the way that reference-price effects can influence the consumer's decision and welfare. Whether it is an appropriate modelling device remains an empirical question.

Axiom 2 *Each consumer is endowed with a price function,* $P(.;.), \mathbb{R}_+^2 \to \mathbb{R}_+$, *which for every reference price,* p^R, *maps from the actual prices,* p, *into adjusted prices,* p^A. *Consumers treat the adjusted price as a sufficient statistic for all price-related information.*

A crucial element of my model is the characterization of the way reference prices distort the actually observed prices. There are some guiding forces towards such a

[4] I am assuming here that the reference price does not affect the intrinsic valuation for the item. This is appropriate when the price does not serve as a signal of quality and when we are thinking of a good for private consumption, not for resale.

formulation. First, if the good is free the reference price should have no effect. Second, the price function should capture the empirically established and sensible idea that the effect of a reference price is to make a price that is higher than it look (even) higher and a price that is lower than it look (even) lower. By the same token, when the observed price coincides with the reference price there should be no distortion. Finally, we would expect the adjusted price to be non-negative. Thus, I would like to posit the following:

Axiom 3 *The price function has the following properties:*

a) $P(0;.) \equiv 0$

b) $P(x;x) = x \ \ \forall x \in \mathbb{R}_+$

c) $0 \leq P(x;y) < x$ *if* $y > x > 0$

d) $P(x;y) > x$ *if* $0 \leq y < x$.

In order to simplify the proofs and to clarify intuition, I also make the following innocuous assumption:

Assumption 1 $P(.;y)$ *is strictly increasing and differentiable, except for a possible kink at* y.

Note that while I allow for it, I do not assume the equivalent of loss aversion (higher directional derivative of $P(p;y)$ at $p = y$ from above than from below y) in this scenario. The jury is still out on how general the existence of loss aversion is, for example in the context of frequently purchased grocery products (c.f. Bell and Lattin, 2000), which is one of my prime applications.

In most of the paper I will only use Axioms 1-3 to derive my results. However, for illustrational purposes it is useful to have a concrete functional form in mind. Preferably one that is easy to estimate and to use in calculations, and one that makes use of Occam's razor.

I would also like to use the concrete functional form for the price function as an opportunity to add a parameter to its description. This parameter is a measure of the *intensity* of the reference price effect. As for the actual functional form, I propose the following:[5]

$$P_\alpha(p; p^R) := p\left(1 + \alpha \frac{p - p^R}{p^R}\right). \tag{1}$$

Here, α (≥ 0) is the intensity parameter. When $\alpha = 0$, there is no distortion, $P_0(p; .) \equiv p$, while the size of distortion, $\left|P_\alpha(p; p^R) - p\right|$, is increasing in α. This is an appealingly simple formula. Perhaps, the only fly in the ointment is that when α exceeds 1, the adjusted price may turn out to be negative (for p small). For tractability's sake, in the rest of the paper I will simply assume that $\alpha \leq 1$, which is equivalent to an upper bound of 2 on the slope of the price function at the reference price.

3 Willingness to pay

The simplest formulation of consumer choice is to look at a single binary purchasing decision in isolation.[6] Then, a consumer is going to buy if the utility of her current wealth is no greater than the utility of the good and of her current wealth minus the price. The price that would make the consumer indifferent between buying or not is defined as her *valuation*, which we will denote by v. Its defining equation is then

$$U(w, 0) = U(w - v, 1), \tag{2}$$

where w denotes wealth, and the second argument of the utility function is an indicator function.

A consumer's valuation of an object is a basic building block for all trade models, whether disaggregate (say, bilateral bargaining) or aggregate (say, a competitive

[5] This price function is convex. If that is considered inappropriate, say because we might expect the reference price effect to decrease for very high prices, a logistic function could be used as an alternative.

[6] The typical example is buying (or not) a durable good.

market). Hence, the reference price effect on valuations can have far-reaching consequences.

With reference prices, the consumer's willingness to pay, denoted by v^R, does no longer coincide with his valuation, defined above as v. Rather, it is implicitly defined by the *adjusted* price which solves (2):

$$U(w,0) = U\left(w - P(v^R; p^R), 1\right). \tag{3}$$

Note that – as utility is increasing in wealth – this means that the adjusted price when the consumer is charged his willingness to pay equals his valuation: $P(v^R; p^R) = v$. This observation has some direct implications.

Proposition 1 *If the reference price exceeds her valuation for an item, then the consumer will be willing to spend more than her valuation on the item. On the other hand, if the reference price is less than her valuation for an item, then the consumer will not be willing to pay as much as her valuation for the item.*

Proof. Given $P(v^R; p^R) = v$, and that the adjusted price is increasing in the price, by Axiom 3 it is immediate that $p^R > (<) v$ implies $v^R > (<) v$. ■

I show in the Appendix that the aggregate effect of this distortion is to make the (residual) demand curve more elastic, and consequently it results in lower prices in monopoly and in imperfect competition (even with differentiated products). Here, I would just like to emphasize that Proposition 1 provides a simple explanation for some important stylized facts: on the one hand, in periods of generalized price reductions (sales) people tend to buy things that they do not really need (that is, they pay for the item more than what it is worth to them) just because it is a "good deal" – in the sense that the price is lower than the reference price which is identified with the pre-sale price; on the other hand, even when people are extracted all their surplus and pay their "willingness to pay" for an item that they value highly – in the sense that their valuation exceeds their reference price – they are usually happy with their purchase, indicating that *ex post* they are not truly indifferent.[7]

This latter point underlines the importance of distinguishing what goes on in the consumer's mind at the time of the purchasing decision and once she has "cooled off". A variant of this idea was first proposed by Thaler (1985):

3.1 Transaction utility

Thaler proposes a decomposition of the change in a consumer's utility as a result of buying an item into two – additive – components. The first one is just the standard change as a result of losing the price and gaining the item – the "acquisition utility" – while the second one is a function of the difference between the price and the reference price, which he calls "transaction utility", $t(p^R - p)$. This latter function is assumed to be increasing, convex below and concave above zero. In other words, "good" surprises give a positive kick, while "bad" ones result in a negative utility shock. He also assumes that the negative shocks are larger and thus the term exhibits loss aversion ($t(p - p^R) < -t(p^R - p)$ for $p > p^R$). As I show below, the transaction utility function can be thought of as a special case of the price function, at least when welfare considerations are not an issue. Indeed, within the context of a single purchasing decision, my approach can be thought of as a (slight) generalization of Thaler's theory.

Formally, in my notation (which also allows for wealth effects), Thaler's set-up defines the customer's willingness to pay as the price which solves

$$U(w, 0) = U(w - p, 1) + t(p^R - p). \tag{4}$$

Implicitly defining $P'(p; p^R)$ as $U\left(w - P'(p; p^R), 1\right) \equiv U(w - p, 1) + t(p^R - p)$, we can reproduce the Thaler set-up (c.f. (3) and (4)) as long as the resulting $P'(p; p^R)$

[7] Such behavior could be behind the experimental results on bargaining, which are usually explained by people having a mixture of "social" and selfish preferences (c.f. Bolton 1991). The reference price may be seen as the "social focal point". One could analyze standard bargaining protocols with players "suffering" from reference price bias and possibly explain bargaining data.

satisfies Axiom 3. Using the fact that utility is strictly increasing in wealth, this can be easily verified – except that the resulting price function does not always yield a non-negative adjusted price.[8] This minor discrepancy is actually the tip of an iceberg, as we will see next.

In terms of the outcome of a single binary choice, whether the reference price affects the utility of the consumer or the perceived price is unimportant. As we will see in the next sections, once we consider the bigger picture of consumer behavior, the conceptual innovation becomes relevant. First, however, I would like to highlight an important issue regarding welfare. Even if the choices made following either of the two approaches coincide, the consumer's utility derived from the purchase will be different. My model implies that after the fact only the acquisition utility matters, while the Thaler approach would take full account of the transaction utility as well.[9] I believe that the reference-price bias is more important at the point of the purchasing decision than the actual share of the transaction utility of total utility. Some time after the purchase the reference-price effect fades away. Thus, the discounted lifetime value of the good is likely to dwarf the transaction utility. However, at the time of trade, the reference-price effect can be very strong. My model allows for incorporating that strong effect without distorting welfare.[10] Of course, the best model is probably in between. I would be happy to include a transaction utility term as well as the price distortion. It is just more didactic to concentrate on the extreme case where the price function replaces transaction utility.

[8] Note, however, that my approach is more general, as the above process cannot be inverted (the resulting transaction utility function will not be a function of $p - p^R$ in general).

[9] That is why sometimes we need a negative price in the price function to capture the aditional utility incorporated in the transaction utility term.

[10] An alternative could be to incorporate psychological effects on utility in case an opportunity to "save" is missed. Such "regret" would appear on the other side of the inequality and thus not affect the overall utility if the purchase was indeed made. It would be difficult to rationalize the corresponding effect for higher than expected prices though.

4 Optimizing the amount to buy

The consumer's decision becomes more interesting when the choice variable is continuous, rather than binary. According to my formulation, there are no degrees of freedom as to how to formalize this and the corresponding optimization problem is straightforwardly defined as

$$x^* = \arg\max_{x \geq 0} U\left(w - p^A x, x\right),$$

where the second argument is no longer an indicator function, rather it signifies the (real) number of units bought.

That is, since it is the unit price that is distorted, the distortion gets scaled up linearly by the amount bought. If we used the Thaler approach, we would have no theoretical indication as to how the transaction utility should depend on the amount of the purchase. As a result, the two approaches are not directly comparable (however, see the next section).

The first-order condition for a maximum is

$$U_2\left(w - p^A x^*, x^*\right) = U_1\left(w - p^A x^*, x^*\right) p^A. \tag{5}$$

The second-order conditions are $U_{11}\left(w - p^A x^*, x^*\right) < 0$, $U_{22}\left(w - p^A x^*, x^*\right) < 0$, and $U_{11}U_{22} > U_{21}U_{12}$. The first two of these, capture the idea of decreasing marginal value both for wealth and for the good and therefore are likely to be satisfied. The third condition is also sensible, as they basically say that the effect on the slope of the utility function of increasing either wealth or the quantity of the good should be no smaller in the same dimension than in the other one. Since I am looking for interesting applications where there is an interior optimum, I assume that the third condition is also satisfied at the solution to (5).

Differentiating both sides of (5) with respect to p^A and rearranging, we obtain

$$\frac{dx^*}{dp^A} = \frac{x^*\left(U_{21} - p^A U_{11}\right) + U_1}{U_{22} - p^A\left(U_{21} + U_{12}\right) + \left(p^A\right)^2 U_{11}}.$$

Therefore, $U_{21}\left(w-p^{A}x^{*},x^{*}\right) \geq p^{A}U_{11}\left(w-p^{A}x^{*},x^{*}\right)$ and $p^{A}U_{12}\left(w-p^{A}x^{*},x^{*}\right) \geq U_{22}\left(w-p^{A}x^{*},x^{*}\right)$ together are sufficient conditions for consumption to be decreasing in the (adjusted) price. These conditions are consistent with the second-order conditions.

When the sufficient conditions hold, we have the intuitive result that a positive (negative) price surprise, $p^{R}-p>(<)0$, increases (decreases) the amount bought relative to the standard model. As a straightforward consequence – holding the actual price constant – the total spend also varies in the same direction. This means that when a good is unusually cheap the consumer actually spends more as a result of the reference-price effect. Similarly, observing an unexpectedly high price the consumer decreases the amount bought so that overall she ends up spending less. These results implicitly also tell us how the amount purchased of the other goods will vary with today's reference-price effect, as the money left over for those purchases will change. This observation leads us to an alternative formulation of the problem:

5 Consumer choice with a soft budget constraint[11]

The analysis in the previous section perhaps requires too much rationality from our consumer, as she is implicitly required to solve a life-cycle problem, which we know empirically that she can/does not (c.f. Thaler, 1990). In order to simplify her problem, let us make the usual assumption about intertemporal separation, namely that she has a certain amount of money to spend in the current period. Thus, we consider a consumer choosing her preferred bundle of n goods given a budget. According to my theory, reference prices affect her choice via the prices and therefore, exclusively through the budget constraint. In effect, given her budget, B, she solves:

[11] Pun intended! It was another Hungarian, János Kornai, who introduced the term *soft budget constraint* in his analysis of a command economy, where budgets were cut if firms underspent and debts written off if firms overspent. His concept has nothing to do with consumer choice though.

$$\arg\max_{\mathbf{x}\geq 0} u\left(x_1, x_2, ..., x_n\right) \quad \text{s.t.} \ \sum_{i=1}^{n} p_i^A x_i \leq B. \tag{6}$$

Note that the crucial consequence of this formulation is that the budget constraint will be satisfied for the *adjusted*, rather than the *actual* prices. As a result, in general the consumer will either under- or overspend relative to her budget![12]

Understandably, this may sound odd at first. Without a saving motive, spending less than the budget is sub-optimal, while spending more than it sounds infeasible. However, it actually makes a lot of sense, if one recalls that the static model is just a sub-problem of a dynamic scenario. My basic assertion is that the amount of money available for consumption at a given point in time is not fixed. The budgets consumers use in their optimization are only 'notional,' they mainly exist to simplify the complexity of the problem. I am not assuming that budgets are fully elastic and income can costlessly be moved across periods. Rather I claim that budgets are not fully inelastic. For most purchase decisions of most consumers (in the developed countries) there is sufficient flexibility – think of a credit card, for example – to borrow against future income (or additional effort!) at a (small) cost. Similarly, there is no major obstacle to saving, even retaining the real value of funds.

The heterogeneity of budget elasticities in the population can be neatly captured by differing intensities of the reference-price effect (α). This brings us to an interesting observation: it may make sense to treat this parameter asymmetrically around the reference price. A person with low borrowing potential may be unaffected by un-

[12] Note that if we added a transaction utility term to the objective function, the chosen bundle would be unaffected (if interior), as the transaction utility only depends on the unit price. If the transaction utility entered the objective function in a multiplicative way – thereby making the transaction utility associated to each good dependent on the quantity purchased of it – it would in general affect the chosen bundle, but the budget would still be exhausted. This could lead to situations where if all goods resulted in a similar sticker shock, the outcome (though not the welfare) would be the same as without the surprise. In my model, the consumer would adjust total spending according to the sign of the shock. I will revisit this issue later in this section.

expectedly low prices (low α), but be still very sensitive to unexpectedly high prices (high α). Note that this would be very similar to the effect of loss aversion.

Even solving problems like (6) are quite difficult for the average consumer, especially when they have lots of items to buy.[13] The natural way boundedly rational consumers simplify their optimization problem is to have notional budgets also for subsets (categories) of the goods they are buying.[14] So much for fruit, for milk products, for wine etc. Then they can separately optimize in each subset. The next subsection formalizes this idea, followed by the formal analysis of the consumer's choice with a soft budget constraint.

5.1 A simple model of consumer behavior

Let the planning horizon of our consumer be T periods. She has some disposable income, M, for the planning horizon, to spend on planned purchases.[15] There are $K \geq T$ categories of goods that she buys. We denote by x_k the amount (or quality[16]) she purchases of each category. Based on her experience and the available information, she forms an expectation of the price for each category, p_k^E. Thus, she can calculate the optimal bundle to buy during the planning horizon using these expected prices (that is, the reference prices). This will result in notional budgets for each period.

From the consumer's point of view prices are random and she only finds them out when she observes them, period by period. As actual prices may differ from the expected ones the consumer may suffer a "sticker shock". The effect of this surprise is a price distortion, which is modelled following Section 2 of this paper. As argued above, this may lead to an outlay which is either above or below the notional budget.

Crucially, however, the imbalance need not affect consumer behavior in the sub-

[13] Think of a weekly grocery shopping, for example.

[14] See Heath and Soll (1996) and Thaler (1999), for a discussion of theory and evidence about the mental budgeting process.

[15] Exceptional expenses are ignored in this version of the model.

[16] In order to interpret x_k as the quality rather than quantity of a product, we would need to consider non-linear prices. While this is clearly doable, for simplicity, I stick to linear prices.

sequent decisions (within the same planning horizon). Note that the consumer does "feel" that she has just spent her notional budget, even if this is not the case. Consequently, until she "counts her money" and re-optimizes[17] – which we assume she is doing every T periods – the amount actually spent in a period will not influence her future behavior. The size of T is thus an important parameter. Crucially, it cannot be infinity (or very large) as that would amount to the consumer ignoring her lifetime budget constraint.

One way of dealing with this problem could be to require that in period T the true budget constraint for the entire planning horizon must be satisfied. This is feasible, if we assume that while the consumer does not count her money until the last period of the planning horizon, she does notice when she runs out of it, say, by reaching her credit limit (or, in a supermarket, she can return some of the items from her shopping cart).

5.2 Optimal choice with a soft budget constraint

I now derive the optimal bundle the consumer chooses within each period. I assume that, for the purpose of this analysis, the groupings of categories into the same period are exogenously given – say, because they are sold in the same shop (or even in the same aisle). Let the consumer have the single-parameter, constant-elasticity-of-substitution (CES) utility function. To simplify matters, I assume that there is only one good subject to reference-price bias and the rest can be grouped together as if they were a single commodity. That is, our consumer solves the following problem

$$\arg\max U(x_1, x_2) = [x_1^\rho + x_2^\rho]^{\frac{1}{\rho}} \text{ s.t. } p_1 x_1 + p_2^A x_2 \leq m.$$

[17] That the frequency with which consumers (or investors) take stock of their activities can be a relevant parameter was first shown by Benartzi and Thaler (1995), where they could explain the equity premium puzzle by using loss aversion and yearly evaluation of stock returns.

It is straightforward to derive the resulting demand functions (for $\rho < 1$):

$$x_1^* = \frac{m p_1^{\frac{1}{\rho-1}}}{p_1^{\frac{\rho}{\rho-1}} + (p_2^A)^{\frac{\rho}{\rho-1}}},\; x_2^* = \frac{m \left(p_2^A\right)^{\frac{1}{\rho-1}}}{p_1^{\frac{\rho}{\rho-1}} + (p_2^A)^{\frac{\rho}{\rho-1}}}.$$

We then have the following result:

Proposition 2 *Relative to the standard model*

1. *Own demand, x_2^*, and own expenditure, $p_2 x_2^*$, are always higher (lower) if the own price, p_2, is below (above) the reference price.*

2. *Cross demand, x_1^*, and cross expenditure, $p_1 x_1^*$, are higher (lower) if either the own price, p_2, is below (above) the reference price and $\rho < 0$, or if the own price is above (below) the reference price and $\rho > 0$.*

3. *For Cobb-Douglas preferences ($\rho = 0$), cross demand and cross expenditure are unaltered by the reference-price effect.*

Proof. The derivative of x_2^* with respect to p_2^A is easily seen to be negative for $\rho < 1$. As $p_2^A > (<) p_2 \Leftrightarrow p_2^R < (>) p_2$, the first result follows. The derivative of x_1^* with respect to p_2^A is easily seen to be negative (positive) for $\rho < (>) 0$. Again $p_2^A > (<) p_2 \Leftrightarrow p_2^R < (>) p_2$ implies the second result. Finally, the derivative of x_1^* with respect to p_2^A is zero for $\rho = 0$. ■

Unlike in the life-cycle model above, now we can compare some of the predictions of Thaler's model with those of mine. While we would expect that the qualitative effects on own demand would be the same with the transaction utility approach, cross demand would necessarily have to move in the opposite direction to compensate for the under- or over-spend. Thus, for relatively complementary goods ($\rho < 0$) the effects on cross demand are in opposing directions for the two approaches. I predict movement in the same direction for both goods, while Thaler's model implies changes in opposing directions. At the end of the spectrum, for Leontieff preferences ($\rho = -\infty$), the transaction utility approach would result in no change from the standard optimal bundle $x_1^* = x_2^* = \frac{m}{p_1+p_2}$, while my model predicts $x_1^* = x_2^* = \frac{m}{p_1+p_2^A}$.

5.3 Precautionary under-budgeting

An interesting question is how a rational consumer, who knew that she suffered from reference-price bias, would try to mitigate its effects.[18] As at the time of purchase she cannot help herself, the only way to influence her decision is by setting "sub-optimal" notional budgets. Note that decreasing the soft budget assigned to a (joint) purchasing decision unambiguously lowers the amount spent (for every realization of the actual price). With Cobb-Douglas preferences the variance of the amount spent also decreases, as the purchase decisions become separable and the notional budget enters the error term multiplicatively: $\left(1 - \frac{p}{p^A}\right) w^n$. Thus, under the reasonable assumption that overspending is more costly than underspending, Cobb-Douglas preferences should lead to conservative notional budget assignments. This is consistent with Pennings et al. (2005) where they find that budget-constrained consumers tend to spend less than their budget.

Note, however, that for other types of preferences the result is not obvious as in the calculation of the variance of spending the "covariances" enter the picture. For example, if there are two goods and the consumer has a very high reference price for one of them but a very low one for the other the expected value of $\left(p_1^A - p_1\right)\left(p_2^A - p_2\right)$ will be negative, which might lead to the overall variance being increasing in the notional budget.

6 Empirical evidence

Heilman et al. (2002) found that consumers given a ($1.00) coupon[19] at the store entrance for an item that they planned to buy tended to increase their spending on unplanned items – relative to a control group – by (seven times) more than their

[18] This is related to the idea that mental budgeting can be used for self-control purposes, c.f. Thaler and Shefrin (1981).

[19] Unfortunately, a coupon is (presumably) only valid for a single unit, so it is not equivalent to a price reduction.

coupon saving. I am somewhat sceptical about the authors' claim that the experimental subjects did not anticipate the windfall ($5.00) after their shopping from the organizers. Also, the results seem to be more driven by "mood" effects, which I – as an economist – am not qualified to appraise. In any case, my model – in its current form – is not apt for predicting unplanned purchases. The spending on planned items remained constant. Unfortunately, the value of the coupon was too low relative to the total spend to make it possible to establish whether the total spending on the planned items was affected by the coupon's value. As a result, the data cannot be used to detect spill-overs to other planned items. What is clear is that these spill-overs could not be large, even if they existed. In conclusion, the findings of Heilman et al. (2002) do not shed much light on the empirical validity of my model's predictions.

Janakiraman et al. (2006) provide a seemingly much better suited experimental design for out purposes. In their main study the subjects are asked to minimize the aggregate cost of stocking a dozen different goods over 35 periods. Their inventory of any good cannot exceed 4. They can go shopping in each period at a fixed cost, but they must shop in case they stock out of something (stocks are depleted "randomly" from the point of view of the subjects). Upon such "mandatory" shopping trips the price of the stocked-out item is raised or lowered by 80%.

They find that there is a negative spill-over effect: if prices are high (low) in a category the consumer spends more (less) on the other categories. Unfortunately, this set-up is not amenable to establish whether my prediction, that consumers increase spending on sale items and decrease spending on unusually expensive items (let alone the resulting spill-overs), holds. This is so because, by construction, one can only spend 80% of the standard budget – 4 items at 20% of the price – on sale items and one must spend at least the price of 1.8 units – 1 unit at 180% of the price – on overpriced items. As the average uptake of regularly-priced units was 1.6, this cannot result in an underspend.

Genesove and Mayer (2001) look at seller, rather than buyer, behavior. However, there is no reason to believe why the price distortions would look qualitatively different from a seller's point of view. Therefore, I can use their analysis to test my model. They look into the explanation of why sellers are setting "unreasonably" high asking prices during economic downturns, that is when they are likely to realize a (nominal) loss relative to the purchase price. They claim that their empirical results validate the loss aversion hypothesis. However, they are only looking at losses, not at gains. Therefore, what they really show is not that losses have higher effects on utility than gains, simply that losses have an effect relative to the standard theory. But that means that their findings are supporting my theory just as well as loss aversion.

7 Conclusion

I this paper I propose a fresh way of looking at reference price effects in a model of consumer choice. The idea that it is the prices, not the preferences that are distorted has interesting consequences, witness the concept of a soft budget constraint.

There are a number of limiting factors of the current model. It relies on marginal effects and as such divisibility of the products is important. It also side-steps dynamic inventory optimization, thereby being more applicable for perishables.

While I have provided a number of testable predictions above, there are lots of implications of my approach that could not be followed up here. I am especially thinking of dynamic pricing and models of negotiation. These could be followed up in another paper.

References

[1] Ariely, D., Kőszegi, B., Mazar, K. and Shampan'er, K. (2004) "Price-Sensitive Preferences," mimeo. UC Berkeley, May.

[2] Beggs, A. and Graddy, K. (2006) "Anchoring effects: Evidence from Art Auctions," mimeo. Oxford, November.

[3] Bell, D.R. and Lattin, J.M. (2000) "Looking for loss aversion in scanner panel data: The confounding effect of price response heterogeneity," *Marketing Science*, 19(2), 185-200.

[4] Benartzi, S. and Thaler, R. (1995) "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110(1), 73-92.

[5] Bolton, G. (1991) "A comparative model of bargaining: Theory and evidence," *American Economic Review*, 81, 1096-1136.

[6] Deaton, A. (1977) "Involuntary Saving Through Unanticipated Inflation," *American Economic Review*, 67(5), 899-910.

[7] Genesove, D. and Mayer, C. (2001) "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 116(4), 1233-1260.

[8] Heath, C. and Soll, J.B. (1996) "Mental Budgeting and Consumer Decisions," *Journal of Consumer Research*, 23(1), 40-52.

[9] Heidhues, P. and Kőszegi, B. (2005) "The Impact of Consumer Loss Aversion on Pricing," mimeo. UC Berkeley, May.

[10] Heidhues, P. and Kőszegi, B. (2006) "Competition and Price Variation when Consumers are Loss Averse," mimeo. UC Berkeley, January.

[11] Heilman, C., Nakamoto, K. and Rao, A. (2002) "Pleasant Surprises: Consumer Response to Unexpected In-Store Coupons," *Journal of Marketing Research*, 39(2), 242-252.

[12] Janakiraman, N., Meyer, R.J. and Morales, A.C. (2006) "Spillover Effects: How Consumers Respond to Unexpected Changes in Price and Quality," *Journal of Consumer Research*, 33(3), 361-369.

[13] Kiyotaki, N., and Moore, J. (2003) "A Cost of Unified Currency." In *Central Banking, Monetary Theory and Practice: Essays in Honor of Charles Goodhart*, Vol. 1, edited by Paul Mizen, 247–255. Cheltenham, UK: Edward Elgar.

[14] Kőszegi, B. and Rabin, M. (2006) "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121(4), 1133-1066.

[15] Mazumdar, T., Raj, S.P. and Sinha, I. (2005) "Reference Price Research: Review and Propositions," *Journal of Marketing*, 69, 84-102.

[16] Pennings, J.M.E., van Ittersum, K. and Wansink, B. (2005) "To spend or not to spend? The effect of budget constraints on estimation processes and spending behavior," *Advances in Consumer Research*, 32, 328-329.

[17] Putler, D.S. (1992) "Incorporating Reference Price Effects into a Theory of Consumer Choice," *Marketing Scienc*e, 11(3), 287-309.

[18] Thaler, R. (1980) "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior and Organization*, 1, 39-60.

[19] Thaler, R. (1985) "Mental Accounting and Consumer Choice," *Marketing Science*, 4(3), 199-214.

[20] Thaler, R. (1990) "Saving, Fungibility and Mental Accounts," *Journal of Economic Perspectives*, 4(1), 193-205.

[21] Thaler, R. (1999) "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12(3), 183-206.

[22] Thaler, R. and Shefrin, H.M. (1981) "An Economic Theory of Self-Control," *Journal of Political Economy*, 39, 392-406.

[23] Tversky, A. and Kahneman, D. (1991) "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *Quarterly Journal of Economics*, 106(4), 1039-1061.

[24] Winer, R. (1986) "A Rference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research*, 13, 250-256.

8 Appendix

8.1 Monopoly

In this section we investigate the effect of the consumers' perception of prices according Axioms 1-3 on the price set by a monopolist.

Assume that in the standard context (no reference price), demand would be given by the differentiable function, $D(p)$, where $D(1) = 0, D(0) = a$ for some $a < \infty$, and $-\infty < D' < 0$. Assume as well that the demand function satisfies the monotone hazard rate property – that is, that $D(.)/D'(.)$ is increasing.

Assuming, for simplicity, that the cost of production is zero, it is immediate that the monopoly price would be the unique interior solution to

$$D(p^M) = -p^M D'(p^M). \tag{7}$$

Now, suppose that the monopolist learns that the way consumers perceive prices is according to Axioms 1-3, with a common reference price and price function for all.[20]

The first issue to look at is how the demand function is transformed as a result of reference price dependence. That is, what is the relationship between the original demand and the one created as a result of reference price dependence. Let us denote the "new" demand curve by $D_R(.)$. Then we have that $D_R(p) \equiv D(p^A(p))$. By Axiom 3, $D_R(p^R) = D(p^R)$, that is, at the reference price the two demand curves coincide. In addition, we also have that $D_R(p) < D(p)$ for $p > p^R$ and $D_R(p) > D(p)$ for $p < p^R$. In other words, $D(p^A(p))$ is qualitatively a counter-clockwise rotation of $D(p)$ around

[20]The commonality of the reference price follows endogenously, see below. The common price function is assumed just for simplicity.

the point $(p^R, D(p^R))$. It is important to note though that, if and only if the price function has slope one at the reference price, will the slopes of $D_R(.)$ and $D(.)$ coincide at $p = p^R$.

What will the new monopoly price be?

We can write total revenue as $pD(p^A)$, yielding the first-order condition

$$D(p^A) = -pD'(p^A)\frac{dp^A}{dp}. \tag{8}$$

Note that the optimal adjusted monopoly price will be a function of the (common) reference price of the consumers.

If it is common knowledge that the monopolist is aware of the reference price dependence of consumer demand, then the consumers rationally[21] expect the monopolist to set a price solving (8). This price will then naturally play the role of the reference price in their adjusted price function.[22] The monopoly price in this Rational Expectations Equilibrium must then be equal to the reference price.

Proposition 3 *The monopoly price under rational expectations is lower than the monopoly price in the absence of a reference price: $p^{RE} \leq p^M$. If the left derivative of $P(p, p^M)$ at $p = p^M$ is > 1, then $p^{RE} < p^M$.*

Proof. If the price function is differentiable, – from (8) – the RE price is given by the unique solution to

$$D(p^{RE}) = -p^{RE}D'(p^{RE})\frac{dP(p, p^{RE})}{dp}[p^{RE}]. \tag{9}$$

By Axiom 3 $\frac{dP(p,p^{RE})}{dp} \geq (>)1$. By the monotone hazard rate property of the demand function, this implies $p^{RE} \leq (<)p^M$. Let us denote the left and right derivatives d_L/d

[21] We could have adaptive expectations as well, but the interesting price is the one where the monopolist settles.

[22] Note that even if the price functions were hetereogeneous, as long as their distribution were common knowledge, the rational expectations price would be common to all consumers.

and d_R/d, respectively. Then (9) can be rewritten as

$$\begin{aligned} D(p^{RE}) &\leq -p^{RE}D'(p^{RE})\frac{d_L P(p,p^{RE})}{dp}[p^{RE}], \\ D(p^{RE}) &\geq -p^{RE}D'(p^{RE})\frac{d_R P(p,p^{RE})}{dp}[p^{RE}]. \end{aligned} \tag{10}$$

Since $\frac{d_i P(p,p^{RE})}{dp} \geq 1$, $i = L, R$, there is clearly no incentive to increase the price beyond the monopoly price. By the same token, a downward change is always profitable if and only if $\frac{d_L P(p,p^{RE})}{dp} > 1$. ■

Putting the technical issues aside, the intuition is clear: as the (inverse) demand function is "rotated" counter-clockwise, the price elasticity of demand is decreased at the original monopoly price, so unit-elasticity is reached at a lower price.

8.2 Oligopoly

In this section we investigate to what extent the price reducing effect observed in the monopoly context carries over to a market with (imperfect)[23] competition. To do this, we will look at a Cournot oligopoly. It is quite intuitive that the effects observed in case of monopoly will also appear when an oligopolist reacts to a residual demand curve. In order to establish this result, we write the total revenue of producer i as pq_i. This leads to the first-order condition

$$\frac{\partial p}{\partial p^A}\frac{\partial p^A}{\partial q_i}q_i + p = 0.$$

Using that $\frac{dD(p^A)}{dq_i} = D'(p^A)\frac{\partial p^A}{dq_i} = 1$, and that $nq_i = D(p^A)$, we obtain the equilibrium condition

$$D(p^{RE})\frac{\partial p}{\partial p^A}[p^{RE}] = -np^{RE}D'(p^{RE}).$$

When p^A is differentiable, by Axiom 3, $\frac{\partial p}{\partial p^A}[p^{RE}] < 1$, while in the standard model it is 1. Thus, we have shown that

[23] It is obvious that under perfect competition there will be no such effect.

Proposition 4 *When the price function is differentiable, the Cournot price under rational expectations is strictly lower than the Cournot price in the absence of a reference price:* $p^{RE}(n) < p^{C}(n)$*.*

The question we would like to shed some light on next, is the interplay between the amount competition and the reference price effects. In order to be able to look at comparative statics, Let us assume that (inverse) demand is linear, given by $p = A - Q$, where Q is the aggregate quantity produced. Also, let us use the specific functional form proposed above, (1).

Proposition 5 *In the rational expectations equilibrium of a Cournot competition with* n *producers*[24] *the equilibrium price is given by*

$$p^{RE}(n,\alpha) = \frac{A}{1+n(1+\alpha)}.$$

Proof. Each competitor maximizes pq_i in q_i.Inverting the price function it is immediate that the actual price as a function of the reference and adjusted prices is given by

$$p = -\frac{(1-\alpha)\,p^R}{2\alpha} + \sqrt{\left(\frac{(1-\alpha)\,p^R}{2\alpha}\right)^2 + \frac{p^R p^A}{\alpha}}. \tag{11}$$

Since consumers use p^A to guide their decisions, the (inverse) demand is given by $p^A = A - \sum q_i$. Taking the first-order condition of pq_i, imposing symmetry and using that in equilibrium $p = p^A = p^R$, we have

$$p(1+\alpha) = q = \frac{A-p}{n}.$$

Solving for p yields the claimed result. ■

In this proposition the effects of the level of competition and the intensity of the reference price effect are confounded. The interesting question is how the *relative* reference price effect varies with n.

Corollary 1 *The proportional reduction in the price as a result of the reference price effect increases with the number of competitors,* $\frac{\partial\left(p^{RE}(n,\alpha)/p^{RE}(n,0)\right)}{\partial n} = -\frac{\alpha}{(1+n(1+\alpha))^2} < 0$.

[24] Note that the proposition also covers the monopoly ($n = 1$) case.

8.3 Horizontally differentiated products

In a Hotelling set-up, for any transportation cost that does not cause technical difficulties, the demand functions are once again more own-price elastic. Thus, for fixed locations, prices will be lower with reference-price effects.

As a further consequence, the increased price competition will tend to increase the incentives for differentiation relative to the standard case.

In a monopolistically competitive market, again prices will be lower. This will alleviate the problem of operating below the efficient scale, but increase the business stealing effect. As a result, the overall effect on the amount of entry is in general indeterminate.

8.4 Vertically differentiated products

In this section I investigate how the presence of reference price effects influences the reaction of consumers of a good provided by a monopolist to the introduction of an inferior substitute by a competitor. This setup is motivated by the wide-spread worry that globalization (understood as international trade) undermines the market viability of locally produced high quality goods (c.f. Kiyotaki and Moore, 2003).

Assume that the incumbent (Firm 1) and the entrant (Firm 2) are located at the two endpoints of a two-unit-long Hotelling interval. Consumers are uniformly distributed between the incumbent and the midpoint of the interval, capturing the vertical differentiation aspect. Travel costs are linear and each consumer values the good at 2 minus his travel cost. Finally, the constant marginal cost of the incumbent is 1, while that of the entrant is zero. For tractability, I also assume that the price function is linear: $p_i^A = p_i + \alpha(p_i - p^R)$.

Proposition 6 *When the entrant affects the market share of the incumbent, the latter is* $\frac{3-\alpha}{6}$.

Proof. The entrant will only affect the market share of the incumbent if all consumers purchase. In that case the marginal consumer is given by the x value solving $x + p_1^A = 2 - x + p_2^A$, leading to $x^*(p_1, p_2) = \frac{2+(1+\alpha)(p_2-p_1)}{2}$. The profits of Firm 1 are given by $(p_1 - 1)\, x^*(p_1, p_2)$, leading to the best response function $p_1(p_2) = \frac{2+(1+\alpha)(p_2+1)}{2(1+\alpha)}$.

Under full market coverage, the entrant's problem is to maximize $p_2\,(1 - x^*(p_1, p_2))$, which leads to $p_2(p_1) = p_1/2$. Otherwise, the entrant's prices do not affect the incumbent's market share.

Solving for the prices with full market coverage, we obtain $p_1 = \frac{6+2\alpha}{3(1+\alpha)}, p_2 = \frac{3+\alpha}{3(1+\alpha)}$. The resulting marginal consumer will be $\frac{3-\alpha}{6}$. ■

Thus the incumbent market share is strictly decreasing in the intensity of reference price dependence, significantly increasing the business-stealing effect.

Note that the market share is independent of the reference price in the above proposition. This is due to the linearity of the price function. If we used the quadratic formulation introduced earlier instead then we would have an additional effect lowering the incumbent's market share as the marginal consumer would be negatively related to the reference price (holding the actual prices constant).

The general idea is quite clear even without a formal analysis. The incumbent firm will always be the one that determines the reference price. Even if the entrant's quality is lower, the reference price will always be above the price the entrant would set in a world without a reference-price effect. As a result, the entrant's prices will always be seen favorably while the incumbent's unfavorably, leading to a larger equilibrium market share for the entrant.