Trade-in transactions typically involve an exchange of an old, used version for a new or newer version of the product. When consumers trade in their used model for a new model, the firm faces the choice of paying the consumer a relatively low price for the used model and charging a commensurately low price for the new model or paying a relatively high price for the used model and charging a commensurately high price for the new model. The extant literature suggests that consumers always prefer to be overpaid in trade-in transactions because they disproportionately value the gain associated with the revenues from the sale of the used version of the product. The authors draw from the prospect theory value function to develop a simple analytical model that identifies a condition under which this preference for overpayment is reversed. Their model predicts that even when faced with economically equivalent price formats, consumers prefer to be overpaid when the ratio of the price of their used product to the price of the new product is low, but when that ratio is high, the preference for overpayment is reversed. They observe support for the predictions that emerge from the model in laboratory experiments.

Keywords: trade-ins, behavioral decision theory, analytical models, consumer behavior, pricing


In many purchase situations, consumers engage in transactions that involve the purchase of a new model of a product and the simultaneous sale of a previously owned version of the product. In the case of consumer durables such as automobiles, golf clubs, wristwatches, and exercise equipment; nondurables such as men’s suits and athletic shoes; and industrial products such as CT scanners, textile machinery, and executive jets, such trade-in transactions are becoming relatively commonplace. In effect, the acquisition price of the new model is reduced by an amount equal to the value of the used model.

Trade-ins are ubiquitous and represent a significant economic activity. Approximately 40% of all new consumer automobile purchases involve a trade-in (Morton, Zettelmeyer, and Silva-Risso 2003). They are also becoming increasingly common in business markets for products such as supercomputers (Fudenberg and Tirole 1998). Furthermore, in developing countries, which have experienced an increase in demand for new durables, trade-ins (often referred to as “exchanges”) are popular (Jain 1993) for a host of products ranging from watches to television sets. Thus, in terms of sheer magnitude, trade-ins represent tens of billions of dollars’ worth of economic activity.

In addition, trade-ins have attracted the attention of policy makers and consumer watchdog groups for several reasons.
On the one hand, trade-in practices have raised concerns about the generation of consumer confusion. For example, in 2003, the attorney general of Illinois accused dealers of engaging in trade-in practices that were “nightmarish challenges for consumers” because of the complexity associated with the deal (Consumer Affairs 2003). Similarly, electronic car buying guides often refer to trade-ins as tools that dealers employ to “rip off” or “confuse” customers by inflating the trade-in price and simultaneously overpricing other elements of the deal, such as the price of the new model, the cost of the warranty, or the magnitude of the interest rate (www.InsiderCarSecrets.com, www.edmunds.com). On the other hand, at times, policy makers have used trade-ins to spur the sale of desirable replacements, such as the sale of new and environmentally friendly cars. For example, in 1995 and 1996, French provincial governments offered consumers subsidies akin to trade-ins to scrap their used cars for new ones (Adda and Cooper 2000). In the United States, the recently concluded “Cash for Clunkers” program, designed to stimulate consumer spending, indicates the efficacy of such trade-in programs. Therefore, in light of the apparent economic and public policy significance of the issue, an understanding of the behavioral underpinnings of the trade-in transaction is a topic worthy of serious analytical and empirical scrutiny.

There are several plausible reasons for the existence of trade-in or joint transactions, rather than the two transactions being treated as separate, individual trades. With trade-ins, consumers can reduce their “hassle” costs because they can sell their used models without engaging in costly advertising, responding to telephone inquiries, providing repeated demonstrations of the product, and so on. Moreover, the analytical literature suggests that trade-ins improve the firm’s ability to (1) remove older models of a product from the marketplace to mitigate their cannibalizing impact on sales of the new model (Levinthal and Purohit 1989), (2) price discriminate (Van Ackere and Reyniers 1993), and (3) mitigate information asymmetry problems regarding older models (Rao, Narasimhan, and John 2009).

Our interest in the trade-in phenomenon is different. We focus on whether the manner in which the firm frames the trade-in transaction has an impact on the preference exhibited by the consumer. We draw from an emerging “behavioral” tradition, which relies on principles of prospect theory and mental accounting, to address whether economically equivalent trade-in deals are psychologically more or less palatable to the buyer (Purohit 1995; Zhu, Chen, and Dasgupta 2008). Specifically, we address whether economically irrelevant differences in the pricing structure of the trade-in transaction yield meaningful differences in consumer preferences.

The extant wisdom in the marketing literature indicates that, in general, consumers prefer to be “overpaid” for the old model they want to sell, even if that overpayment results in a commensurately higher price for the new model they want to purchase (Purohit 1995; Zhu, Chen, and Dasgupta 2008). We develop an analytical model and offer empirical evidence that identifies the conditions under which this preference for overpayment may be reversed and consumers will exhibit a preference for an “underpayment” for the model they want to sell. Our prediction is that when the ratio of the value of the old model to that of the new model reaches a critical point, consumers’ preference for overpayment on the old model reverses. That is, they would rather be underpaid on the old model with a concomitant reduction in the price of the new model when the ratio of the value of the old model to that of the new model is higher than this critical point. We empirically observe this preference reversal in a series of experimental studies.

The contribution of our research is fourfold. First, consistent with recent treatments of the topic, we demonstrate that consumers are not indifferent between economically equivalent alternatives but rather display a theoretically defensible preference for over- and underpayments for their previously owned models. Second, we offer a simple analytical model as well as empirical evidence to support our claim. Third, our findings are of potential practical significance for sellers of durable products who want to frame the trade-in transaction and to policy makers who want to foster trade-ins for certain welfare-improving interventions. Fourth, our study sheds new theoretical insights on the rapidly growing field of behavioral pricing, specifically in the area of partitioned prices. Prior research has argued that consumers respond more favorably when they encounter prices that have been partitioned into constituent elements (e.g., the price of the product and shipping and handling fees) than when separate pricing elements are combined (Chakravarti et al. 2002; Morwitz, Greenleaf, and Johnson 1998). Our inquiry on trade-ins focuses on transactions whose prices are naturally partitioned into two elements: the price associated with the old model and the price associated with the new model.

BACKGROUND LITERATURE AND MODEL

An important property of trade-in transactions relevant to our theorizing is the availability of two pieces of price information: one associated with the old or trade-in model and one associated with the new model. Therefore, before we turn to the theoretical argument that motivates our model and empirical test, we briefly review relevant literature that examines settings in which consumers encounter multiple prices for a single purchase transaction, or the literature on partitioned prices.

Partitioned Prices

The literature on partitioned prices (e.g., Morwitz, Greenleaf, and Johnson 1998) demonstrates that when a product’s total price can be partitioned into two separate elements (e.g., the product’s price and shipping and handling), because the magnitude of one salient constituent element (the product’s price) is lower, demand may increase. Chakravarti and colleagues (2002) extend the theory and observe that, consistent with mental accounting principles, when evaluating prices associated with multicomponent bundles, consumers prefer partitioned prices. Furthermore, they observe that the effects are moderated by the attention paid to the price of the focal product. That is, “decision outcomes vary with different splits of the same total bundle price” (Chakravarti et al. 2002, p. 227, emphasis added).

Other studies in this area have highlighted the consequences of partitioning. For example, Hamilton and Srivastava (2008) find that in a bundled product, when the higher consumption value component is priced higher, the overall evaluation of the price increases. Cheema and Soman
(2008) highlight the role of partitioning product quantities and find that consumption is regulated (declines) when products (e.g., chocolates) are partitioned into multiple units compared with when they are aggregated.

In our study, a trade-in transaction can be thought of as a bundled transaction with two components: a new model being purchased and an old model being sold. The important difference between trade-in transactions and the partitioning of prices (or products) is that in trade-in transactions, consumers purchase one product (in which case they incur an expenditure and lose money) while selling another (in which case a monetary benefit accrues to them and they gain money), whereas in the partitioning of prices, both elements of prices represent expenditures. This distinction allows for an analysis of how consumers might process the prices associated with the trades and the development of some theoretically testable predictions regarding the price formats consumers would most prefer.

A Simple Model

We employ Kahneman and Tversky’s (1979) prospect theory value function to separate the trade-in transaction into two components. One component reflects the gain associated with the revenue that accrues to the consumer from selling the old model, and the other component reflects the loss associated with the payment the consumer incurs when purchasing the new model.

The prospect theory value function is characterized by three properties (Chen and Rao 2002; Kahneman and Tversky 1979; Thaler 1985). First, outcomes are coded separately for gains and losses in comparison with some reference point (typically a person’s current state of wealth), implying that (1) people respond to changes relative to the reference point, not absolute states, and (2) the two outcomes are additively separable. These two premises of the first property are important for our model development. Specifically, as Thaler (1999, p. 188) observes, the notion that “costs are generally viewed as losses” has been rejected (Kahneman and Tversky 1984; Thaler 1985). Losses and gains are outcomes relative to a consumer’s current state of wealth. Therefore, we need to ensure that respondents behave in a manner consistent with our model’s first premise. With respect to the second premise, if the outcomes are not separable and are combined into one outcome, consumers will simply evaluate the overall arithmetic consequence of the difference in the two outcomes. We empirically examine both these premises in a pretest associated with Study 1 and in Study 2.

Second, a property of the value function is that it is S shaped: concave for gains and convex for losses (i.e., \( v''(x) < 0, x > 0 \); \( v''(x) > 0, x < 0 \)). This property reflects the psychophysical principle of diminishing marginal sensitivity.

Third, the response is steeper for losses than for gains (i.e., \( v(x) < -v(-x), x > 0 \)), reflecting the principle of loss aversion, according to which the disutility associated with a loss is greater than the utility associated with a gain of the same magnitude. Using the prospect theory value function, we first present the intuition that underlies the results we obtain and then present a formal model from which the results flow.

Model Intuition

Consider two trade-in settings. In one setting, the value of the model to be traded in is low, and in a second setting, it is high, relative to the value of the model being purchased. First, consider the low relative value condition. For example, the consumer may be trading in a used cell phone for a new one. Then, we can compare preferences for two economically equivalent pricing formats. In one case, the vendor might offer a trade-in deal comprising an allowance of $40 for the trade-in model (a gain relative to the status quo ante) and a price of $250 for the new instrument (a loss relative to the status quo ante). In another case, the vendor might offer a trade-in deal comprising an allowance of $80 for the trade-in model (a gain) and a price of $290 for the new instrument (a loss). When comparing the two (economically equivalent) options, consumers may not be indifferent. The lower values associated with the gains ($40 and $80, which are the prices associated with model to be traded in) are closer to the origin, and because people are relatively sensitive to changes close to the origin since the slope of the value function close to the origin is relatively steep, they should be relatively sensitive to the difference in magnitudes of the gains. Conversely, the higher values associated with the losses (prices associated with the new model) are further from the origin, and because the slope of the value function far from the origin is relatively shallow, people should be relatively insensitive to the difference in magnitudes of the losses. Therefore, because the difference between the two gains ($40 and $80) has a greater impact on value perceptions than the difference between the two losses ($250 and $290), people should prefer the higher gain ($80) and be less affected by the accompanying higher loss (Purohit 1995; Zhu, Chen, and Dasgupta 2008). Figure 1 presents this intuition.

In the second setting, the value of the model to be traded in is high relative to the value of the model being purchased. For example, consider again a consumer trading in a used cell phone. The vendor might offer several economically equivalent trade-in deals: a trade-in allowance of $210 for

![Figure 1](image-url)

**Figure 1**

**VALUATION OF A TRADE-IN TRANSACTION WHEN THE USED MODEL’S PRICE IS LOW RELATIVE TO THAT OF THE NEW MODEL**

\[
\begin{align*}
V(–$250) + V($40) & = V(–$290) + V($80) \\
V($40) & > V($80) \\
V(–$250) & > V(–$290)
\end{align*}
\]
the trade-in model (a gain) and a price of $250 for the new instrument (a loss) or a trade-in allowance of $250 for the trade-in model (a gain) and a price of $290 for the new instrument (a loss). Now, the values associated with the gains (prices associated with the model being traded in) are far from the origin, and people should be relatively insensitive to changes in gains. The higher values associated with the losses (prices of the new model) are also far from the origin, and people should be relatively insensitive to changes in losses as well. However, because the prices associated with the new model are situated in the loss domain, the sensitivity to losses may be greater because of loss aversion. As a consequence, the preference for a relatively high price for the trade-in model may be reversed, and people may prefer a relatively low price on the new model (and the accompanying low price for the trade-in model). This preference reversal is a consequence of the interplay of two forces: the curvature of the value function, which yields differences in slope close to and far from the origin, and the relative strength of the loss aversion coefficient in the domain of losses. Figure 2 presents this intuition. Next, we develop a mathematical model that allows for a more formal description of this interplay.

The Model

A trade-in transaction consists of the following four key elements: acquisition (purchase) of a new model, trade-in (sale) of a used model, the price associated with the new model, and the receipt of a “trade-in allowance” for the used model. We employ these elements in the construction of our model and assume that (1) consumers evaluate each element separately, (2) the value of each element enters additively into the overall subjective valuation, (3) the quality of the products is fixed and cannot be changed by either party (i.e., there are no moral hazard problems), (4) the valuation of each element follows the principles of prospect theory’s value function, and (5) individual consumers face uncertainty about prices in the marketplace so that their current state of wealth is the reference price. In other words, in the trade-in transaction, the entire monetary receipt of the trade-in offer is coded as a gain because monetary receipts are desirable, while the entire monetary payout for the new item is coded as a loss because monetary outflows are undesirable and aversive.1

We define the following model terms:

\[ \eta = \text{valuation of quality of a new model of the product;} \]
\[ \mu = \text{valuation of quality of the used model being traded in;} \]
\[ p_n = \text{stand-alone price of a new model, offered during a transaction that does not involve any trade-ins;} \]
\[ p_u = \text{stand-alone price that a dealer will offer to pay for the used model being traded in;} \]
\[ p_n > p_u > 0. \]

As we discussed previously and will be evident subsequently, the quality-related parameters are irrelevant for our derivations. We present them here for conceptual completeness.

The reference-dependent value function can be expressed as follows (Kivetz 2003):

\[
(1) \quad v(x) = \begin{cases} 
 g(x) & \text{if } x \geq 0 \\
 (-\lambda)g(-x) & \text{if } x < 0
\end{cases},
\]

where \( g(0) = 0, g' > 0, g'' < 0, g''' \geq 0, \) and \( \lambda > 1. \)

All our results can be derived from the general functional form given in Equation 1. Here, for expositional convenience and to enable us to focus on the trade-off between loss aversion and diminishing sensitivity, we use a simple and tractable functional form that captures all the key properties of the value function. Note that a functional form that includes different curvatures in the domains of gains and losses (i.e., \( x^{\alpha} \) under gains and \( x^{\beta} \) under losses; Chen and Rao 2002; Tversky and Kahneman 1992) is also amenable to this analysis. However, as Tversky and Kahneman (1992) empirically observe, \( \alpha = \beta; \) thus, for mathematical tractability, we employ the same exponent under losses and gains.2 Therefore, we consider the following:

\[
(2) \quad v(x) = \begin{cases} 
 x^{\alpha} & \text{if } x \geq 0 \\
 (-\lambda)(-x)^{\alpha} & \text{if } x < 0
\end{cases},
\]

where \( \lambda > 1 \) and \( 0 < \alpha < 1. \) Loss aversion is captured with \( \lambda, \) and \( \alpha \) captures the diminishing sensitivity of the value function.

Mathematically, we can write the subjective value of a trade-in transaction in which a retailer quotes a new model price of \( p_n \) and offers a trade-in allowance of \( p_u \) for the used model being traded in using Equation 2 as follows:

\[
(3) \quad v_{\text{td}} = \eta^{\alpha} - \lambda(\mu)^{\alpha} + p_u^{\alpha} - (\lambda)p_n^{\alpha} = \left[ \eta^{\alpha} - \lambda(\mu)^{\alpha} \right] - \left[ (\lambda)p_n^{\alpha} - p_u^{\alpha} \right].
\]

1To the extent that a transaction can be decoupled into “get” (product attributes and benefits) and “give” (monetary sacrifice) components, it is essential for our model that monetary payments are deemed aversive (loss inducing). People should prefer to pay less rather than more because paying is painful, and paying more is even more painful. Recent evidence in consumer behavior (see Rick, Cryder, and Loewenstein 2008) and cognitive neuroscience (see Knutson et al. 2007) is consistent with this premise.

2For formal derivations based on the general functional form in Equation 1, see Appendix B. For formal derivations based on a functional form that incorporates different exponents in the domains of gains and losses, see the Web Appendix at http://www.marketingpower.com/jmrfeb11.
Suppose that the retailer quotes a price \((p_n + d)\) for a new model and offers a price \((p_u + d)\) for a used model in a trade-in transaction, where \((-p_u) < d < (p_n - p_u)\). Because \((p_u + d) - (p_u + d) = p_n - p_u\), the retailer should be indifferent between these two price formats.

However, from the perspective of the consumer engaging in the trade-in, the transaction that involves prices \((p_u + d)\) and \((p_n + d)\) yields a total subjective value of the following:

\[
(4) \quad v_{00} = \frac{\eta^\alpha - \lambda(\mu)^\alpha + (p_u + d)^\alpha - (\lambda)(p_n + d)^\alpha}{(\lambda)(p_n + d)^\alpha - (p_u + d)^\alpha}.
\]

A comparison of Equations 3 and 4 indicates that the two valuations are not necessarily the same, even though the economic consequences to the consumer are the same. Next, we turn to a derivation of a proposition and three corollaries that flow from this observation. (Appendix A provides all the proofs.)

**Proposition**

Our key proposition is intended to identify the conditions under which consumers prefer an overpayment or an underpayment on their trade-in option. To accomplish this derivation, we define two terms: \(p_u/p_n \equiv k\) represents the ratio of the price of the used model to the new model and lies between 0 and 1 by construction, and \(Q \equiv \lambda^{1/\alpha - 1}\), a transformation of the loss aversion coefficient, which also lies between 0 and 1 because \(\lambda \geq 1\) and \(0 < \alpha < 1\). This term captures the interplay between the curvature of the value function (\(\alpha\)) and the loss aversion coefficient (\(\lambda\)). The following proposition shows that a consumer’s most preferred pricing format in a trade-in transaction depends on \(k\):

\(P_1\): The optimal deviation of pricing from stand-alone prices is \(d^* = [(Q - k)/1 - Q]p_u\). This deviation is decreasing in \(k\), positive for \(k < Q\), and negative for \(k > Q\). In other words, as the ratio of the price of the trade-in model to the price of the new model increases, the consumer experiences a preference reversal from desiring an overpayment on the trade-in allowance to desiring an underpayment on the trade-in allowance when evaluating two financially equivalent transactions.

**Proof:** The consumer’s utility is maximized for a value of \(d\) satisfying argmax \([\{n\alpha - \lambda(\mu)^\alpha\} - \{(\lambda)(p_n + d)^\alpha - (p_u + d)^\alpha\}]\), implying \(d^* = [(Q - k)/1 - Q]p_u\).

In other words, consumers prefer to be overpaid on the model being traded in when the value of the used model is relatively low compared with that of the new model, and they prefer to be underpaid on the model being traded in when the value of the used model is relatively high compared with that of the new model.

Figure 3 is a graphical representation of this result, in which loss aversion (\(\lambda\)) appears on the horizontal axis and \(Q\) (the tipping point that determines the reversal) appears on the vertical axis, for three different levels of \(\alpha\). In this figure, a pair of \((\lambda, \alpha)\) values yield a unique \(Q\), such that for \(k < Q\), overpayment is preferred, and vice versa. Figure 3 shows the regions for the preference of over- and underpayment.

Another straightforward implication of \(P_1\) is that for \(k > 1\) (i.e., when consumers are “trading down”—when the value of the used item is higher than the new item being purchased), \(d^* < 0\) (i.e., underpayments are always preferred; for a formal proof, see Appendix A). We capture the intuition underlying this result in the following simple example: Imagine a consumer engaging in a trade-in transaction in which she is selling an object worth \$20,000 (she enjoys a monetary gain) while buying an object worth \$10,000 (she suffers a monetary loss). The impact of underpayment on \$10,000 leading to a reduction in loss is substantially higher than the impact of overpayment on \$20,000 leading to an increase in gain because of the relative proximity of \$10,000 to the origin and because of loss aversion. Thus, when this consumer is trading down, both diminishing sensitivity and loss aversion yield a preference for an underpayment.

Next, we present three corollaries of this result. The first two corollaries demonstrate that both loss aversion and diminishing sensitivity acting together are necessary to observe a reversal in preference from a preference for overpayment to a preference for underpayment. In the third corollary, we demonstrate that the cutoff point \(Q\) changes with changes in the loss aversion coefficient \(\lambda\).

**Corollary 1:** In the absence of loss aversion (\(\lambda \to 1\)) and the presence of diminishing sensitivity (\(\alpha < 1\)), consumers always prefer an overpaid trade-in allowance to an underpaid trade-in allowance between two financially equivalent transactions.

Figure 3 indicates this result. The cutoff point \(Q\) is obtained by examining any \((\lambda, \alpha)\) pair. For any value of \(\alpha < 1\) and \(\lambda = 1\), \(Q = 1\). For underpayment, \(k\) must be greater than \(Q\), but because \(k < 1\) while \(Q = 1\), in this case, no region exists in which underpayment is preferred.

**Corollary 2:** In the absence of diminishing sensitivity (\(\alpha = 1\)) and the presence of loss aversion (\(\lambda > 1\)), among financially equivalent transactions, consumers’
most preferred transaction involves an underpaid trade-in allowance.

Again, Figure 3 shows the intuition for this result. As we move from lower to higher values of $\alpha$, the region of overpayment ($k < Q$) shrinks. In the limit, as $\alpha \to 1$, the overpayment region shrinks to zero. Figure 4 further illuminates the intuition. In this graph, we plot the optimal allowance as a function of $k$ for different values of $\alpha$, with the loss aversion parameter fixed at 2.0. The point at which a line intersects the x-axis determines the value of $k$ above which underpayments are preferred. As the sensitivity decreases (i.e., $\alpha$ increases), the point at which the reversal occurs becomes smaller, and in the limit, $\alpha \to 1$, $k = 0$, which implies that underpayments are preferred for the entire range of $k$.

Together, Corollaries 1 and 2 demonstrate the fundamental trade-off that a consumer faces. Diminishing sensitivity leads to a preference for an overpaid pricing format, while loss aversion leads to a preference for an underpaid pricing format. Intuitively, when the value of the used model is a small fraction of the price of the new model, the extent to which even a small addition to the gain from the sale of the used model payment is pleasurable is greater than the extent to which the same small increase to the loss from the payment made for the new model is aversive. In other words, when the values associated with the gain are small, the diminishing sensitivity effect dominates the loss aversion effect, and therefore consumers prefer the overpayment pricing format. Conversely, when the value of the used model is a large proportion of the value of the new model, a small reduction in the loss associated with the payment for the new model is more valuable than a commensurate decrease in the gain from the payment received from the used model. In this case, the loss-aversion effect dominates the diminishing sensitivity effect, and consumers prefer the underpayment pricing format. In essence, these two corollaries highlight the opposing effects of loss aversion and diminishing sensitivity and provide a rationale for why the predicted reversal occurs. The tipping point for the predicted preference reversal is reflected in $Q$, and Corollary 3 shows how this tipping point $Q$ changes with changes in $\lambda$.

Corollary 3: For any given value of $\alpha$, the point $Q$ that determines the cutoff for preference reversal from overpayment to underpayment on trade-in allowance is a function of the loss aversion coefficient ($\lambda$). Specifically, the larger the loss aversion coefficient, the larger is the range of $k$ observed for consumers’ preferences for underpayment in trade-in allowances. That is, the higher the value of $\lambda$, the lower is the value of $k$ that generates a negative $d^*$. Figure 3 reveals the intuition that underlies this result. If we examine the graph for any $\alpha$, as we move to the right (increasing $\lambda$), $Q$ declines. In other words, for higher values of the loss aversion parameter, the preference reversal is obtained for lower values of $k$. This result becomes clearer in Figure 5, in which we plot the optimal allowance as a function of $k$ for different values of $\lambda$ (1.5, 2.0, 2.5, and 3.0, respectively) while keeping $\alpha$ fixed. The point at which a line intersects the x-axis determines the value of $k$ above which underpayment will be preferred. As Figure 5 shows, for higher values of $\lambda$, the intersection occurs at a lower value of $x$, indicating a larger range in which underpayment is preferred. Next, we describe our empirical efforts to assess support for our claims.

**EMPIRICAL TESTING**

All the studies we present employ an experimental approach in which participants were exposed to stimuli for a consumer durable available from a hypothetical store. The scenario to which consumers were exposed described an opportunity to exchange an existing used product for a new one. At one store, the used model to be traded in was to receive a relatively low price (e.g., $40) and the new model’s sticker price was commensurately low (e.g., $250), and at the other store, the used model to be traded in was to receive a relatively high price (e.g., $80) and the new model’s sticker price was commensurately high (e.g., $290). The total cash outlay required of participants was identical (e.g., $210) at both stores.
Depending on the purpose of the study, we employed a series of dependent variables designed to elicit store preferences. In addition, we measured theoretically justifiable individual difference variables and manipulation checks. In all, we report three principal studies. In the first, we assess support for our foundational prediction that a preference reversal will be observed depending on the ratio of the used model trade-in price to the new model price. In the second study, we confirm that this effect occurs only when the transactions are coded separately; when the transactions are combined, the effect disappears. In the third study, we examine a prediction that emerges from Corollary 3, according to which the magnitude of loss aversion should have an impact on the preference reversal.

**STUDY 1**

The purpose of Study 1 is to test the proposition derived in the preceding section, which we restate as a refutable prediction here:

\( H_1: \) When trading in a model whose value is a relatively small fraction of the price of the new model being purchased, consumers will prefer to be overpaid on the model they are trading in while being overcharged by a commensurately higher amount for the new model being purchased, relative to an economically equivalent option that underpays on the model being traded in and charges a commensurately lower price for the new model. Conversely, when trading in a model whose value is a relatively large fraction of the price of the new model being purchased, consumers will prefer to be underpaid on the model they are trading in while being charged a commensurately lower amount for the new model being purchased, relative to an economically equivalent option that overpays on the model being traded in and charged a commensurately higher price for the new model.

Recall that two important premises underlying our model are that (1) cash outlays are coded as losses and cash receipts are coded as gains, relative to a reference point (a person’s current state of wealth), and (2) outcomes are additively separable. We assess the validity of the first premise using a pretest here and the validity of the second premise experimentally in Study 2.

**Pretest**

Fifty-two undergraduate students at a large U.S. university participated in a short in-class survey. In one condition, they were asked to imagine purchasing a digital camera outright, priced at $250, and in another condition, they were asked to imagine trading in their old camera, for which they would receive a trade-in value of $250. In each condition, participants responded to two dependent measures designed to elicit their perceptions regarding the dollar amount being expended on the new camera or received for the trade-in (1 = “gain compared with current status,” and 7 = “loss compared with current status”). The pretest results confirm that participants perceived the price paid on the new camera in the outright purchase condition as a loss to a significantly greater degree than the price received for the old camera in the trade-in condition (4.35 versus 2.46; t(51) = 4.84, p < .01). In other words, the money received for the trade-in model is indeed viewed as a gain relative to the participant’s current wealth state, while the purchase price of the new product is viewed as a loss, relative to the participant’s current wealth state. Thus, an essential assumption of our model is tenable, which allows us to proceed with our empirical inquiry.

**Design**

To examine empirically whether our foundational prediction regarding the anticipated preference reversal would be observed, our principal focus in Study 1 was on a manipulation of the ratio of the price of the used model to the price of the new model. We measured respondents’ preferences for the two pricing formats, which we embedded in the stimulus. Therefore, our first task was to design a stimulus in which the ratio of the trade-in model’s price to the new model’s price could be manipulated. In particular, we wanted to employ prices of the trade-in model that were either relatively low or relatively high, in comparison with the price of the new model, while holding all other variables constant.

To accomplish this task, we created four price combinations for the stimulus product (cameras). In one price combination (low ratio), we priced the trade-in model at $40 and the new model at $250, and in the second, we priced the trade-in model at $80 and the new model at $290. In the third price combination (high ratio), we priced the trade-in model at $210 and the new model at $250, and in the last price combination, we priced the trade-in model at $250 and the new model at $290. This set of price combinations represents the two levels of one factor in our experimental design.\(^4\)

We also manipulated a second factor. Here, we provided additional reference price information about the trade-in model, the new model, or both. Our reason for providing this information was to ensure empirically that respondents were indeed coding the trade-in amount as a gain (not as a loss) and coding the price of the new model as a loss (not as a gain), relative to some reference price. We accomplished this manipulation by providing a suggested price for the new model (e.g., “You know that the suggested price of this camera is $320”) and/or the purchase price of the old model (e.g., “which you had paid $100 [or $300] for”). Note that the reference price for the new model ($320) is always higher than the actual selling price, allowing for the possibility of the price of the new model to be coded as a gain (of +$70 when the price is $250 and +$30 when the price is $290) and thus allowing for a fair test of whether participants were coding the price of the new model as a loss. Moreover, and for a similar reason, the reference price for the used model is always higher than the trade-in price ($100 when the trade-in price is either $40 or $80, and $300 when the trade-in price is either $210 or $250), allowing for the possibility of the receipt from the trade-in model to be coded as a loss (of -$60 or -$20, when the trade-in prices are $40 and $80,

\[^4\]Our choice of price stimuli approximately corresponds to the parameter values \(x = 2.25\) and \(a = .5\). These values correspond approximately to the mean of the values reported in the empirical studies of the value function parameters. Although there is broad agreement about the estimates of \(x\), some of the best-known studies show substantial variation in the estimates of \(a\). Tversky and Kahneman (1992) estimate \(a = .88\), Camerer and Ho (1994) report \(a = .32\), and Wu and Gonzalez (1996) report an estimate close to .88. Our experiment implies a value of \(a\) closer to the mean of these three studies. For further discussion on prospect theory parameterizations, see Neilon and Stowe (2002).
respectively, and \(-$90\) and \(-$50\), when the trade-in prices are \$210 and \$250, respectively). We should not observe the predictions of our model unless the participants code receipts associated with the trade-in model as gains and the costs associated with the new model’s price as losses.\(^5\) This procedure yielded a 2 (ratio of trade-in price to new model’s price: low versus high) × 3 (reference price information: for used model, for new model, or for both) between-subjects factorial design.

**Participants and Stimuli**

Fifty-five undergraduate students (approximately 50% female) from a large U.S. university participated in the study. We used a digital camera as the stimulus. Participants were randomly assigned to one of the six experimental cells and read stimuli associated with that cell. For example, respondents assigned to the low ratio and both reference prices provided condition read the following scenario:

You go to some electronic stores to buy a particular digital camera that you want (the PowerShot SX300). You know that the suggested price of this camera is \$320.

Two electronic stores offer trade-ins for an old camera (the EasyPhoto ER50), which you had paid \$100 for.

One store, XX, offers you \$40 on the trade-in for the old camera (the EasyPhoto ER50) and quotes \$250 as the price of the new digital camera (the PowerShot SX300).

The other store, YY, offers you \$80 on the trade-in for the old camera (the EasyPhoto ER50) and quotes \$290 as the price of the new digital camera (the PowerShot SX300).

In both stores, you need to pay \$210 plus the old camera in order to buy the new camera.

Participants assigned to the high-ratio conditions read a similar scenario, though the trade-in allowances were different: \$210 for store XX, and \$250 for store YY.

After they read the scenarios, participants reported their preferences for each store on a ten-point scale (\(-5\) = “Buying at store XX is better than buying at store YY,” and \(+5\) = “Buying at store YY is better than buying at store XX,” with no midpoint, to ensure that participants would be forced to select one or the other option). Subsequently, this scale was transformed to a 1–10 scale. After participants responded to preference items and demographic measures, we debriefed, thanked, and dismissed them.

\(^5\)If participants code trade-in values as losses relative to the original price paid (the relevant reference point) and new model prices as gains relative to the prevailing market price, when the trade-in value is low, the experienced loss is large, but so is the experienced gain. Conversely, when the trade-in value is high, the experienced loss is small, but so is the gain. Because the slope of the value function is steeper close to the origin, and because of loss aversion, the net impact of a small loss and small gain (\(-$20, +$30\)) is likely to be more aversive than the net impact of a large loss and a commensurately large gain (\(-$60, +$70\)). Therefore, if participants code trade-in receipts as losses and the new model’s price as a gain, they should prefer the overpayment format (which yields \(-$60, +$70\)) because it is less aversive than the underpayment format. If we observe a main effect caused by the provision of this additional reference price information, it would imply that, contrary to our model’s assumptions, participants code trade-in receipts as losses and new model prices as gains.

**Results**

We predicted that the store employing the overpayment strategy would be preferred in the low-ratio condition and that the store employing the underpayment strategy would be preferred in the high-ratio condition. Moreover, we expected no main or interaction effects caused by the reference price information manipulation, because we expected respondents to code trade-in receipts as a gain and the price of the new model as a loss, relative to the status quo ante. A 2 × 3 analysis of variance yielded a nonsignificant interaction between reference price information and the ratio of prices (F(2, 49) < 1). The main effect of the presence of reference price information was also not significant (F(1, 49) = 2.02, p < .10). These results support our premise that trade-in receipts are not coded as losses and the price of a new model is not coded as a gain.

As predicted, we observed a significant main effect of the ratio of the price of the trade-in model to that of the new model (F(1, 49) = 5.72, p < .02). Participants preferred the store with the overpayment pricing format when the ratio of the trade-in model to the new model was low (M = 5.64) relative to when the ratio of the trade-in model to the new model was high (M = 4.20).

**A Rival Explanation**

A rival explanation for our results is that in the low-ratio condition, the two trade-in prices ($40 and $80) are relatively salient because they are noticeably different from the retail prices associated with the relevant trades ($250 and $290), while in the high-ratio condition, the trade-in prices ($210 and $250) are less salient because they are almost indistinguishable from the retail prices associated with the relevant trades ($250 and $290). Thus, in the high-ratio condition, participants’ attention may have been focused on the retail price to a relatively greater degree. To assess whether this possibility accounts for our results, we conducted a study (n = 31) at a large U.S. university.

Students participated in an in-class survey, during which they were exposed to the low-ratio condition ($40/$250, $80/$290) and were asked whether they attended to $40 price ($80 price) for the old camera more than $250 ($290) for the new camera on a seven-point scale (1 = “I attended to $40 [$80] price more,” and 7 = “I attended to $250 [$290] price more”). In both the stores, the respondents’ attention to the used camera’s price was not significantly different from their attention to the new camera’s price (M$_{40/250}$ = 4.26 versus 4 [neutral value]; t(30) = 1.137, p > .25; M$_{80/290}$ = 4.03 versus 4 [neutral value]; t(30) = .105, p > .90). In addition, to reconfirm that respondents viewed payments as aversive (and coded as losses) and receipts as desirable (and coded as gains), we asked them to indicate on a seven-point scale how they thought of the money received (price paid) for the old camera (new camera) (1 = “I think of these receipts [price] as a gain,” and 7 = “I think of these payments [price] as a loss”). The results confirm that respondents coded the trade-in receipt as a gain and the price paid for the new model as a loss (2.55 < 5.23; t(30) = 6.83, p < .001).

These results not only confirm our original premise regarding the coding of monetary payments as losses and receipts as gains but also rule out the possibility that low statistical power limited our ability to detect an effect for the
rival, attention-based explanation. Our sample size was sufficiently large enough to detect the effect of loss/gain coding; thus, we cannot attribute the absence of an effect for the relative attention hypothesis to lack of power.

Having established support for our foundational prediction, we now turn to Study 2, in which we examine two methodological nuances. Specifically, we assess whether our results are (1) sensitive to the requirement that outcomes are additively separable and (2) not sensitive to the numerical values we elected to manipulate.

**STUDY 2**

A central premise underlying our prediction is that people code each element of the transaction separately. In particular, if the trade-in model’s price and the new model’s price are combined into one outcome, when faced with economically equivalent outcomes, people should be indifferent between stores that underpay versus stores that overpay. The predicted preference reversal should only be observed when people separate the two outcomes. We designed this study to address this issue.

The study employed a 2 × 2 between-subjects experimental design. One factor manipulated whether the transactions were framed as separate or combined, and the other factor manipulated the magnitude of the ratio between the trade-in model’s price and the new model’s price (low versus high). Student participants evaluated the attractiveness of purchase of a digital camera from two fictitious stores.

**Design**

We manipulated whether participants’ focus on the two transactions should be viewed as separate or combined by employing a priming task. Before being exposed to the stimulus regarding camera prices and store selection, participants were randomly assigned to one of the two framing conditions and asked to evaluate Consumer Reports tips for people planning to buy a new car. Participants in the separate condition evaluated the following:

Negotiate one thing at a time.

You’ll get the best deal by keeping the purchase and trade-in negotiations separate. If you allow salespeople to mix the two, it gives them too much of an opportunity to manipulate the deals so that a good price in one area could be canceled out by a poor price in the other. We suggest that you nail down the price of the car you want to purchase first, then discuss your trade-in allowance. Always evaluate the price of a new car and trade-in allowance separately.

Participants in the combined condition evaluated the following:

Negotiate many things at the same time.

You’ll get the best deal by combining the purchase and trade-in negotiations. If you allow salespeople to keep the two separate, it gives them too much of an opportunity to manipulate the deals so that a good price in one area could be canceled out by a poor price in the other. We suggest that you simultaneously nail down both the price of the car you want to purchase and the trade-in allowance. Always evaluate the price of a new car and trade-in allowance together.

After evaluating these primes, participants evaluated two transactions for a camera purchase, as in Study 1.

A second methodological nuance involved varying the price of the new model to achieve the low- versus high-ratio manipulation. As in Study 1, we created four price combinations for the stimulus product (cameras). In the first, in which the ratio of the used model’s price to the new model’s price was low, we priced the trade-in model at $40 and the new model at $250; in the second price combination, we priced the trade-in model at $80 and the new model at $290. In the third price combination, in which the ratio of the used model’s price to the new model’s price was high, we priced the trade-in model at $40 and the new model at $50; in the last price combination, we priced the trade-in model at $80 and the new model at $90. Note that we produced the high ratio by keeping the price of the trade-in model low while lowering the price of the new model, unlike in Study 1.

**Participants and Stimuli**

Ninety-four undergraduate students (approximately 60% female) from two U.S. universities participated in the study. We used a digital camera as a stimulus. We randomly assigned participants to one of the four experimental cells and gave them stimuli associated with each cell.

Participants first read and evaluated Consumer Reports tips, which served to manipulate separation of the two components of the transaction. Next, they read a scenario about the trade-in transaction offered by two fictitious stores: “Calyypo,” which underpaid on the trade-in and charged a commensurately low price on the new model, and “Advantage,” which overpaid on the trade-in and charged a commensurately high price on the new model. In addition, we measured perceptions of the attractiveness of the offers and the stores employing seven-point scales (1 = “Calyypo is much better than Advantage/I would prefer to shop at Calyypo,” and 7 = “Advantage is much better than Calyypo/I would prefer to shop at Advantage”). Then, to check the success of the separation versus combined manipulation, we asked participants to rate their perception of the transaction (1 = “The two dollar amounts were separate,” and 7 = “The two dollar amounts were combined”). The key dependent variable of interest was a discrete choice item that prompted participants to indicate which store they would patronize, assuming they wanted to engage in the transaction. Participants could indicate Calyypo, Advantage, or indifference. After participants completed the questionnaires, we thanked, debriefed, and dismissed them.

**Results**

Participants in the separate prime condition indicated that they considered the dollar amounts associated with the two transactions separate to a greater degree than participants in the combined prime condition (5.19 > 5.87; one-tailed t(93) = 1.84, p < .05). We deemed our manipulation of the frame (separate versus combined) successful.

Our principal dependent variable of interest is store choice. We ran a fully saturated logit model on store choice as a function of frame (A), ratio (B), and their interaction (A × B) on the participants who displayed a store preference. As

---

6In this study, we rely on participants’ coding monetary payments as losses. In Study 3, we manipulate participants’ loss aversion and test whether they behave in a manner consistent with the predictions of our model.
expected, we observed a significant interaction (b = 3.75, p < .05). Consistent with our prediction, when we primed participants to treat the transactions as separate, the preference for overpayment on the trade-in model in the low-ratio condition was substantial (78.6%), while the preference for underpayment was large in the high-ratio condition (87.5%). A test of proportions indicated that this difference was statistically significant (p < .05). Conversely, when we primed participants to consider the transactions in combination, the preference for overpayment in the two conditions (33.3% and 45.6%, respectively) was not statistically different (p > .10).

The results from the analysis of the attractiveness of the store and of the offers were consistent with the results from the discrete choice analysis. Offer attractiveness displayed a significant two-way interaction between frame (separate versus combined) and ratio (low versus high) (p < .0001) that was driven by a preference reversal. In the separate condition, low ratios yielded a preference for overpayment compared with high ratios (4.68 versus 3.29; one-tailed t(90) = 4.58, p < .001), while in the combined condition, we observed the opposite effect (3.56 versus 4.16; one-tailed t(90) = 1.79, p < .05). Store attractiveness displayed a marginal interaction effect (p < .08), driven by a simple effect for overpayment under low ratios in the separate condition (4.58 versus 3.24; one-tailed t(90) = 4.46, p < .001), and no effect under the combined condition (3.81 versus 3.68, p > .35). These results make sense because the magnitude of the effect should be stronger when evaluating the offer than when evaluating the store overall.

Discussion

In summary, the results of this study replicated Study 1 and our foundational prediction, according to which the preference for under- and overpayment is a function of the ratio of the price of the used model to that of the new model only when people treat the two transactions as separate. These results are evident from the analysis of the discrete choice data and the more powerful analysis of the scalar dependent variables. In addition, we eliminated the possibility that our results were an artifact of the particular numerical combination employed, because the high-ratio condition in this study comprised relatively low prices for the used model. In other words, it is less likely that the sheer numerical magnitude of the transaction in the high-ratio condition makes the price more salient, leading to enhanced loss aversion for the larger purchase, which in turn results in a preference for underpayment.

The finding that our results hinge on whether elements of the transaction are coded separately is an important finding that speaks to the generalizability of our model. If consumers do not spontaneously separate elements of the transaction as we have postulated, our results would not be observed. In other words, the external validity of our results is crucially dependent on the degree to which consumers parse trade-in transaction prices, a topic to which we return in the “Discussion” section. We now turn to our final study, in which we assess support for a prediction derived from our model (Corollary 3), according to which the preference reversal phenomenon that H1 predicts should be more pronounced when the loss aversion coefficient is high.

STUDY 3

Our goal in Study 3 is to assess whether participants who differ in the degree to which they are naturally loss averse display the predicted preference reversal to differing degrees. This is a theoretically defensible boundary condition because, if loss aversion is one underlying driver of the effect, to the extent that people are less loss averse, the effect should be attenuated.

Previous research has identified many possible bases for individual differences in loss aversion. For example, according to one research stream (Wang 1996), loss aversion varies by cultural context. Another recent research stream examines a person’s “regulatory focus” as the basis for differences in loss aversion. In light of its wide applicability in the consumer behavior literature in marketing (e.g., Zhu and Meyers-Levy 2007), we examine whether and how a person’s regulatory focus affects their loss aversion and, thus, the tendency to prefer over- rather than underpayments in trade-in transactions.

Regulatory Focus Theory

Research on regulatory focus theory suggests that people differ in their self-regulation system, distinguishing between promotion and prevention focuses (Higgins 1997; Shah and Higgins 2001). People who are promotion oriented are concerned with maximizing positive outcomes, and people who are prevention oriented are concerned with minimizing negative outcomes. These orientations regulate their preferences and behaviors in many ways. Most pertinent to our research is the finding that people who are promotion oriented are particularly sensitive to potential gains, while people who are prevention oriented are particularly sensitive to potential losses, resulting in different loss aversion patterns between the two types (Chernev 2004). That is, because prevention-oriented consumers are concerned with minimizing losses while promotion-oriented consumers are concerned with maximizing gains, loss aversion is likely to be greater for consumers who are prevention oriented than for those who are promotion oriented. Furthermore, such differences suggest that the degree of preference reversal offered in H1 should differ between promotion- and prevention-focused consumers. Specifically, because prevention-oriented people are relatively more loss sensitive in general, their preference for underpayment rather than overpayment should be greater than any such preference differences exhibited by promotion-oriented people. However, recall that diminishing sensitivity also plays a role in whether consumers prefer underpayment; in the high-ratio condition, the effect of loss aversion should dominate the effect of diminishing sensitivity, and in the low-ratio condition, the effect of diminishing sensitivity should be the dominant force. Consequently, prevention-oriented consumers’ preference for underpayment should be most readily apparent in the high-ratio condition. We conducted a study to assess whether these preferences would be observed.

Participants and Design

Because Studies 1 and 2 established that the predictions of our model are observed only when prices are encoded separately, we primed all participants in this study to process the prices separately using the prime described in Study 2. To manipulate the person’s regulatory focus (pro-
motion versus prevention), we employed a procedure that prior research (e.g., Higgins 1997) has successfully used. Participants were first asked to write down either their goals and hopes (promotion condition) or their duties and obligations (prevention condition), in an attempt to prime one or the other regulatory focus.

One hundred eighty-eight undergraduate students in a large university participated in the study. After their regulatory focus had been primed through the writing exercise, they were exposed to economically equivalent over- and underpayment formats. This procedure yielded a 2 (regulatory focus: promotion-oriented versus prevention-oriented) × 2 (ratio of trade-in price to new model’s price: low versus high) between-subjects factorial design.

Other than the regulatory focus inducing writing task, the stimulus material was similar to that used in Study 1. We asked participants to report their preferences for two stores selling cameras using different trade-in and new model prices on a seven-point scale, in which lower values reflected a preference for underpayment. After participants completed the survey, we thanked, debriefed, and dismissed them.

Results

The results of the study yielded a pattern that was broadly consistent with our theorizing. We observed a main effect for promotion versus prevention (M{	ext{promotion}} = 4.53 > M{	ext{prevention}} = 3.96; F(1, 184) = 9.10, p < .005). Contrast analyses indicate that the simple effect of price ratio on store preference was significant for promotion-oriented participants (M{	ext{low-ratio_promotion}} = 4.95 versus M{	ext{high-ratio_promotion}} = 4.16; F(1, 184) = 9.74, p < .005). Furthermore, regardless of ratio condition, participants always preferred overpayment (the means are greater than 4.0), and the pattern of the means was consistent with Corollary 1 because participants preferred overpayment more when the ratio of prices was low (condition in which the effect of diminishing sensitivity was relatively large). Finally, the simple effect of price ratio on store preference was significant for prevention-oriented participants as well (M{	ext{low-ratio_prevention}} = 4.52 versus M{	ext{high-ratio_prevention}} = 3.49; F(1, 184) = 15.95, p < .001), and the means indicate that, consistent with Corollary 2, such prevention-oriented participants prefer underpayment in the high-ratio condition because the effect of loss aversion dominates the effect of diminishing sensitivity, whereas they prefer overpayment in the low-ratio condition. In other words, for prevention-oriented participants alone, we observed a preference reversal, consistent with our theorizing.

In summary, the results of Study 3 provide a boundary condition for the preference reversal that is consistent with our model. When loss aversion is not substantial (because of a promotion- rather than prevention-orientation), the preference reversal we observe is less likely to manifest.

**DISCUSSION**

**Summary**

We develop a model and conduct experiments that demonstrate a reversal in preference for pricing formats in trade-in transactions, depending on the ratio of the price of the used model to that of the new model. When we separated elements of the transaction, participants displayed a systematic preference for being overpaid on their used model when the ratio of the prices was low, and this preference was reversed when the ratio of prices was high, a finding consistent with our model, which we derived from the prospect theory value function.

**Implications**

Theoretically, and consistent with other recent research, we demonstrate that consumers are not indifferent between economically equivalent pricing formats when evaluating trade-in transactions. Although we observe support for a preference for overpayment, as Zhu, Chen, and Dasgupta (2008) and Purohit (1995) demonstrate, we predict and observe a theoretically defensible reversal of this preference caused by loss aversion and marginal sensitivity. Our simple model, which is based on prospect theory’s value function, allows for the development of relatively precise predictions and provides a parsimonious explanation for the observed preferences. Specifically, the model enables us to specify the levels of the independent variable (see footnote 4) and to generate predictions regarding the role of prevention and promotion (using the corollaries). A less formal, verbal theory would likely not have yielded the same level of precision and parsimony.

In addition, our findings are of considerable practical significance. Vendors of durable products ranging from automobiles and skis to computers and golf clubs face the task of persuading consumers to purchase new models while there still is considerable value in the consumer’s existing model. We recognize that the general problem of persuading consumers to trade in used models is an independent (and significant) issue (Okada 2001); however, our question revolves around the framing of the prices associated with the transaction after the consumer has made the decision to engage in a trade. In other words, our emphasis is on the framing of price information and not on the framing of product information, which we take to be fixed. For products such as late-model automobiles, for which, in general, the ratio of the price of the old to the new version is high, vendors would be well advised to select the underpayment format. Conversely, for products such as cell phones, for which, in general, the ratio of the price of the old to the new version is low, vendors would be well advised to select the overpayment format. However, an important caveat is in order with regard to the generalizability of our results: Our results are crucially dependent on consumers’ spontaneously separating the price elements of the trade-in transaction or being forced to separate these price components by the manner in which the information is presented. In the absence of separability, our results would not be observed.

**Limitations and Further Research**

First, our model and experimental settings only account for circumstances in which the price of the old model is less than that of the new model. However, the new model is occasionally cheaper than the old model. Our model is readily applicable to this case, but empirical validation of such a case is a topic left to further research.

Second, our model only holds when trade-in receipts are coded as gains (and not losses) and the new model’s price is coded as a loss (and not a gain). If situations exist in which the trade-in receipt is coded as a loss (e.g., because a consumer’s internal reference price is relatively high) or the
new model’s price is coded as a gain (e.g., because the consumer’s reference price is relatively high), the results of the model will likely not hold. In addition, it is essential for our model that payments be aversive. One procedure to test whether this coding scheme can be interfered with and thus change our results would be to provide credible and salient stimulus-based reference price information. Then, a high reference price for the trade-in model and/or a very high reference price for the new model might yield different results.

The third conceptual issue is the source of loss aversion. The issue of why people tend to be loss averse has attracted much discussion and debate in the behavioral decision theory literature. Some scholars have argued for an evolutionary basis for loss aversion (Wang 1996), and others have suggested that the observed effects are due to experimental artifacts (Cosmides and Tooby 1996). We are agnostic on this issue, though we observe a nexus between loss aversion and the negativity bias literature (Ahuwalia 2002). For our model to work, it is only necessary for loss aversion to exist.

This research demonstrates that consumers treat economically equivalent prices differently in trade-in transactions. A question that emerges from our research is whether firms can earn higher effective profits by carefully combining the price of the new model and trade-in allowance for the used model. Formal models that involve an extension of our simple model could examine questions about firm strategy under monopoly and under competition in light of this consumer bias.

In addition, we view the two transactions as simultaneous rather than sequential. However, it is possible that if the first transaction is the purchase and the second transaction is the trade-in sale, the consumer’s reference price will shift because of the first transaction, and this reference price shift will have implications for the second transaction (Chen and Rao 2002). Similarly, if the trade-in transaction is first and the purchase of the new model is second, again, the reference price may shift. We do not consider these dynamics in our model.

Finally, we present study participants with economically equivalent options in a within-subjects design. The behavioral literature suggests that when consumers are confronted with choices that are equally (un)attractive, but for different reasons, they experience trade-off aversion (Hedgcock and Rao 2009). Such trade-off aversion has been studied using cognitive neuroscientific methods, and perhaps trade-off evaluations that involve implicit trade-offs (gains and losses) could be studied similarly, while employing between-subjects designs.

In summary, we demonstrate that consumers prefer to be overpaid on their used trade-in model only when the ratio of the price between the trade-in model and the new model is low. When this ratio is high, consumers may prefer to be underpaid on their trade-in model, a finding that is rife with theoretical and practical implications.

APPENDIX A: DETAILS OF THE PROOFS

Details of the Proof of Proposition 1

\[
\text{arg} \max_d \{\ln^\alpha - \lambda(\mu^\alpha)\}(\lambda)(p_n + d)^\alpha - (p_n + d)^\alpha].
\]

Because the quality parameters are exogenous to the model, we can write the preceding as follows:

(A1) \[
\arg\max_d (\lambda)(p_n + d)^\alpha - (p_n + d)^\alpha.
\]

The first-order condition yields

(A2) \[
\lambda \alpha (p_n + d)^\alpha - 1 = \alpha (k p_n + d)^\alpha - 1.
\]

Rearranging the terms in Equation A2 yields

\[
(k p_n + d)^\alpha - 1/(p_n + d)^\alpha - 1 = \lambda.
\]

This implies the following:

(A3) \[
(k p_n + d)/(p_n + d) = \lambda^{1/\alpha} - 1.
\]

Note that \[Q \equiv \lambda^{1/\alpha} - 1;\] therefore, Equation A3 can be rewritten as follows:

(A4) \[
(k p_n + d)/(p_n + d) = Q.
\]

Solving Equation A4 for \[d\] yields

(A5) \[
d^* = \{(Q - k)/(1 - Q)\}p_n.
\]

To check for the second-order condition, we followed this process: The second derivative of Equation A1 with respect to \[d\] is

\[
\begin{align*}
\lambda \alpha (p_n - 1)(p_n + d)^{\alpha - 2} + \alpha (\alpha - 1)(k p_n + d)^{\alpha - 2} \\
+ \alpha (\alpha - 1)(k p_n + d)^{\alpha - 2} = 0.
\end{align*}
\]

Inserting Equation A5 in Equation A6 to find the optimality condition for \[d\] yields

\[
\begin{align*}
\lambda \alpha (p_n - 1)[(Q - k)/(1 - Q)]^{\alpha - 2}(p_n)^{\alpha - 2} + \\
+ \alpha (\alpha - 1)(k + [(Q - k)/(1 - Q)])^{\alpha - 2}(p_n)^{\alpha - 2} = 0.
\end{align*}
\]

Taking \[p_n\] out of the parentheses from the preceding equation produces the following:

\[
\begin{align*}
\lambda \alpha (p_n - 1)[(1 + [(Q - k)/(1 - Q)])^{\alpha - 2}(p_n)^{\alpha - 2} + \\
+ \alpha (\alpha - 1)[k + [(Q - k)/(1 - Q)])^{\alpha - 2}(p_n)^{\alpha - 2} = 0.
\end{align*}
\]

On further simplification, this yields

\[
\begin{align*}
\lambda \alpha (p_n - 1)[((1 - k)/(1 - Q))]^{\alpha - 2}(p_n)^{\alpha - 2} + \\
+ \alpha (\alpha - 1)[(Q - Qk)/(1 - Q)]^{\alpha - 2}(p_n)^{\alpha - 2} = 0.
\end{align*}
\]

Pulling out all the common terms:

\[
\begin{align*}
\lambda (p_n)^{\alpha - 2}[1 - k/(1 - Q)]^{\alpha - 2}[\lambda - (Q)^{\alpha - 2}].
\end{align*}
\]

Recall that \[Q \equiv \lambda^{1/\alpha} - 1.\] Note that the terms outside the final brackets are all positive because \[\alpha < 1, k < 1,\] and \[Q < 1.\]

Next, substitute for \[Q\] in the preceding expression only within the final brackets:

\[
\alpha (1 - \lambda)(p_n)^{\alpha - 2}[1 - k/(1 - Q)]^{\alpha - 2}[\lambda - (\lambda^{1/\alpha - 1})(\alpha - 2)].
\]

That equals

\[
\begin{align*}
\alpha (1 - \lambda)(p_n)^{\alpha - 2}[1 - k/(1 - Q)]^{\alpha - 2}[\lambda - (\lambda^{1/\alpha - 1})(\alpha - 2)].
\end{align*}
\]

Note that the expression inside the parentheses is negative, because \[(\alpha - 2)/(\alpha - 1) > 1,\] and this implies \[\lambda(\alpha - 2)/(\alpha - 1) > \lambda\] because \[\lambda > 1.\] So \[\lambda(\alpha - 2)/(\alpha - 1) < 0,\] and therefore Equation A7 is negative. So at the optimal \[d,\] obtained from the first-order condition, we obtain a negative second-order condition, and thus the original expression must be a maxima.
Proof of Corollary 1
With no loss aversion, \( \lambda = 1 \), so the consumer’s valuation for the pricing format is
\[-(p_n + d)^\alpha + (p_n + d)^\alpha.\]
The first derivative of the preceding expression with respect to \( d \) is
\[-\alpha(p_n + d)^{\alpha - 1} + \alpha(kp_n + d)^{\alpha - 1},\]
or equivalently
\[\alpha \left[ -\frac{1}{(p_n + d)^{1 - \alpha}} + \frac{1}{(kp_n + d)^{1 - \alpha}} \right].\]
This expression is always positive because \((kp_n + d) < (p_n + d)\). In other words, as \( d \) increases, the subjective value increases; therefore, consumers would prefer a larger \( d \) to a smaller \( d \) and overpayment on the trade-in allowance to an underpayment on the trade-in allowance.

Proof of Corollary 2
With no diminishing sensitivity, the \( \alpha \) term has a value of 1. Consider the following inequality, which captures the argument that underpayment is preferred: For every \( d^* > 0 \),
\[(\lambda)(p_n + d^*) + (p_n + d^*) < -(\lambda)(p_n - d^*) + (p_n - d^*).\]
This implies that
\[-\lambda p_n - \lambda d^* + p_n + d^* < -\lambda p_n + \lambda d^* + p_n - d^*.\]
After canceling the common terms, we have
\[-\lambda d^* + d^* < +d^* - d^* \rightarrow 2d^* < 2\lambda d^*.\]
Because \( d^* > 0 \), this implies that \( \lambda > 1 \), which is true by definition. Therefore Equation A8 must be true.

Proof of Corollary 3
Recall that \( Q = \lambda^{1/\alpha - 1} \).
\[
\partial Q / \partial \lambda = \left[ \frac{1}{(\alpha - 1)} \lambda^{2 - \alpha} \right] < 0,
\]
because \( \alpha < 1 \) and \( \lambda > 1 \). Therefore, as \( \lambda \) increases, \( Q \) declines.

Now, recall from \( P_1 \) that
\[d^* = \left( \frac{Q - k}{1 - Q} \right) p_n.\]
From this expression, it follows that a decline in \( Q \) leads to a decline in the cutoff value of \( k \) that generates the reversal, indicating that an increase in \( \lambda \) generates preference reversals at a lower value of \( k \).

If the used item is valued higher than the new item \((k > 1)\), the optimal deviation from benchmark level, \( d^* \), is negative. In other words, when consumers “trade down,” they prefer to be underpaid on the trade-in allowance.

Proof: Recall that the optimal deviation from the benchmark is given by the following (from Equation A5):
\[d^* = \left( \frac{Q - k}{1 - Q} \right) p_n.\]
This is always negative for \( k > 1 \) because \( Q < 1 \).

**APPENDIX B: PROOFS USING A NONPARAMETRIC FORM OF THE VALUE FUNCTION**

**P1:** For \( k < q \), consumers prefer a pricing format that yields \( d > 0 \), whereas for \( k \geq q \), consumers prefer a pricing format that yields \( d \leq 0 \). In other words, as the ratio of the price of the trade-in model to the price of the new model increases, there is a preference reversal from desiring an overpayment to desiring an underpayment.

Proof: The reference dependent value function is given by the following:
\[v(x) = \begin{cases} g(x) & \text{if } x \geq 0 \\ -\lambda g(-x) & \text{if } x < 0 \end{cases} \]
\[g(0) = 0, g' > 0, g'' < 0, g'' > 0, \lambda > 1.\]

Using Equation A1, we can write the total value of a joint transaction that involves an allowance \( d \) as follows:
\[\psi(d, p_n) = -\lambda g(p_n + d) + g(kp_n + d).\]
We complete the proof in the four steps detailed next.

**Step 1: Calculating the Optimal Allowance**
Consumers’ most preferred pricing format for a given price is as follows:
\[\arg \max_d -\lambda g(p_n + d) + g(kp_n + d).\]
We obtain the optimal allowance \( d_k \) from the solution of the following first-order condition:
\[-\lambda g'(p_n + d_k) + g'(kp_n + d_k) = 0\]
or
\[g'(kp_n + d_k) = \lambda.\]
We assume that the function \( g \) is well behaved in the sense that an interior solution exists. This holds as long as we assume that the following second-order condition is satisfied:
\[-\lambda g''(p_n + d_k) + g''(kp_n + d_k) < 0.\]

**Step 2: Showing That the Optimal Allowance Is Positive Around the Lower Bound of \( k \) and Negative Around the Upper Bound of \( k \)**
Next, we show that the optimal allowance is positive around the lower bound of \( k \) and negative around the upper bound of \( k \). Note that \( k \in (0, 1) \), which means that the price of the used model being traded in is strictly less than the new model.

**Part a:** Let the optimal allowance be \( d_{k_0} \) when \( k_0 p_n = \varepsilon \), where \( k = k_0 \) is very close to zero. We want to show that \( d_{k_0} > 0 \). We show this by contradiction. Let us suppose that the optimal allowance is an underallowance given by \( -k^+ \), where \( k^+ > 0 \). Because \(-k^+ \) is optimal, it necessarily implies that
\[-\lambda g(p_n - k^+) + g(kp_n - k^+) > -\lambda g(p_n + k^+) + g(kp_n + k^+)\]
or
\[-\lambda g(p_n - k^+) + g(e - k^+) > -\lambda g(p_n + k^+) + g(e + k^+).\]
This implies that
\[(B5) \quad \lambda > \frac{g(\varepsilon + k^{+}) - g(\varepsilon - k^{+})}{g(p_{n} + k^{+}) - g(p_{n} - k^{+})}.\]

Note that \(\lambda\) has a finite upper bound (in most empirical work, the value is close to 2). In contrast, as \(k^{+}\) gets closer to zero, the denominator of Equation B5 approaches 0 (note that because of diminishing sensitivity, the numerator would always exceed the denominator), and as the left-hand side becomes larger, there must be \(\varepsilon^{*}\) such that for \(\varepsilon < \varepsilon^{*}\), Equation B4 is no longer satisfied. Thus, for \(k_{0}\) very close to zero, Equation B4 will be violated, and this implies that the optimal allowance must be an overallowance. In other words, \(d_{k_{0}}\) must be positive.

**Part b:** Let the optimal allowance when \(k = k_{1}\), where \(k_{1}\) is very close to 1, be given by \(d_{k_{1}}\). Let \(k_{1}p_{n}\) be equal to \(p_{n} - \varepsilon_{1}\), where \(\varepsilon_{1} > 0\) is close to zero. Equation B3 yields the following:
\[(B6) \quad \frac{g^{'}(p_{n} - \varepsilon_{1} + d_{k_{1}})}{g^{'}(p_{n} + d_{k_{1}})} = \lambda.\]

Now if the \(d_{k_{1}} > 0\), as \(\varepsilon_{1} \to 0\), the expression in Equation B6 would approach 1 and can no longer be satisfied. Suppose we take \(d_{k_{1}} = -p_{n} + \varepsilon_{2}\), where \(\varepsilon_{2} > \varepsilon_{1}\) and both \(\varepsilon_{1}\) and \(\varepsilon_{2}\) are in the vicinity of zero. Then, we can rewrite Equation B6 as
\[(B7) \quad \frac{g^{'}(\varepsilon_{2} - \varepsilon_{1})}{g^{'}(\varepsilon_{1})} = \lambda.\]

Note that in Equation B7, we can always pick an appropriate \(\varepsilon_{2}\) for a given \(\varepsilon_{1}\) so that the optimality condition is met. Thus, in the preceding equation, we show that when \(k\) is close to its lower bound 0, the optimal allowance is an overpayment, and when \(k\) is close to its upper bound 1, the optimal allowance is an underpayment.

**Step 3: Showing That the Optimal Allowance Is a Decreasing Function of \(k\)**

Note that at a given \(k\), the optimal allowance \(d_{k}\) is given by the following:
\[\frac{g^{'}(kp_{n} + d_{k})}{g^{'}(p_{n} + d_{k})} = \lambda.\]

Using envelope theorem, we have
\[(B8) \quad \partial d_{k} / \partial k = \frac{-p_{n}g^{''}(kp_{n} + d_{k})}{-\lambda g^{''}(p_{n} + d_{k}) + g^{''}(kp_{n} + d_{k})}.\]

Note that the numerator is positive because \(g^{''} < 0\) (due to diminishing sensitivity) and the denominator is negative due to Equation B3. Thus, for \(k \in (0, 1)\), we have the following:
\[(B9) \quad \partial d_{k} / \partial k < 0.\]
In other words, as \(k\) increases, the optimal allowance decreases.

**Step 4**

There is a \(k = q\) such that for \(k \in (0, q), d_{k} \geq 0\) and for \(k \in (q, 1), d_{k} < 0\). Furthermore, this value of \(q\) is unique and is given by the following implicit equation:
\[\frac{g^{'}(qp_{n})}{g^{'}(p_{n})} = \lambda.\]

Note that in Step 2 of the proof, we show that when \(k\) is close to zero, the optimal allowance given by Equation B3 is greater than zero, while for values of \(k\) close to 1, the optimal allowance is less than zero. It is immediately implied by the intermediate value theorem that there must be a \(k\) denoted by \(q\) where the optimal allowance is zero. This value of \(q\) is obtained by substituting \(d_{k} = 0\) in Equation B3, which yields the following:
\[(B10) \quad \frac{g^{'}(qp_{n})}{g^{'}(p_{n})} = \lambda.\]

Uniqueness is shown by contradiction. Suppose that there is another \(k = (k^{*} + \delta)\) where \(\delta > 0\) such that
\[(B11) \quad \frac{g^{'}[(q + \delta)p_{n}]}{g^{'}(p_{n})} = \lambda.\]

Denominators in both Equations B10 and B11 are the same and because \(g^{'} > 0, g^{''} < 0\), which means that \(g^{'}[(q + \delta)p_{n}] < g^{'}(kp_{n})\). However, this implies that Equation B10 can never be true. The proof for \(\delta < 0\) follows similarly.

**Corollary 1:** When consumers exhibit only diminishing sensitivity and no loss aversion, they always prefer an overpaid trade-in allowance.

**Proof:** Note that when \(\lambda = 1\),
\[(B12) \quad u(d, p_{n}) = -g(p_{n} + d) + g(kp_{n} + d).\]

The first derivative of the preceding equation with respect to \(d\) gives
\[-g^{'}(p_{n} + d) + g^{'}(kp_{n} + d).\]

Diminishing sensitivity ensures that the preceding expression is strictly positive, and thus the higher the value of \(d\), the higher is the value of Equation B12. Therefore, consumers always prefer overpayment.

**Corollary 2:** When consumers exhibit only loss aversion and no diminishing sensitivity, they always prefer an underpaid trade-in allowance.

This is true if, for every \(\Delta^{+} > 0\), we have
\[-\lambda g(p_{n} - \Delta^{+}) + g(kp_{n} - \Delta^{+}) > -\lambda g(p_{n} + \Delta^{+}) + g(kp_{n} + \Delta^{+})\]
or
\[\lambda[g(p_{n} + \Delta^{+}) - g(p_{n} - \Delta^{+})] > g(kp_{n} + \Delta^{+}) - g(kp_{n} - \Delta^{+}).\]

If diminishing sensitivity is not present, then
\[[g(p_{n} + \Delta^{+}) - g(p_{n} - \Delta^{+})] = g(kp_{n} + \Delta^{+}) - g(kp_{n} - \Delta^{+}).\]

This immediately implies that \(\lambda > 1\). This is true, by definition.

**Corollary 3:** Preference reversal increases in \(\lambda\). In other words, the larger the loss aversion coefficient, consumers’ preference for underpayment in trade-in allowance is observed for a larger range of \(k\).
Proof: We show the change in $k^*$ with respect to $\lambda$, by applying envelope theorem to Equation B10, which yields

$$\frac{\partial q}{\partial \lambda} = \frac{g'(p_n)}{p_n g'(kp_n)}.$$ 

It immediately follows from Equation B1 that this is negative. Therefore, for stronger loss aversion, the cutoff $q$ is smaller, which implies that the preference reversal occurs earlier.

REFERENCES