Realization Utility

with Reference-Dependent Preferences†

Jonathan E. Ingersoll, Jr.
Yale School of Management

Lawrence J. Jin
Yale School of Management

†We thank our colleagues at the Yale School of Management and the seminar participants at the University of Massachusetts, Amherst for helpful discussions. In particular we would like to thank Nick Barberis. Jin acknowledges support from a Whitebox Advisors grant.
Realization utility, the concept that investors derive their utility only from actively realizing gains or losses, has drawn increasing interest recently. In this paper we develop a model of realization utility that significantly extends existing models both psychologically and institutionally. We admit either standard risk averse or S-shaped prospect theory utility functions. We use a specification that allows the reference point to be constant, growing or random; it can also depend on recent price history as well as the purchase price. Transactions costs and taxes can affect investors’ subjective views of gains and losses in various ways. These features provide possible explications of a number of aspects of the data that were not possible with earlier models. In particular, loss realization can be optimal, though it is generally much less common than the realization of gains.

JEL classification: G02, G11, G12
Keywords: Realization Utility, Prospect Theory, Disposition Effect
Realization Utility
with Reference-Dependent Preferences

1. Introduction

How do investors evaluate their own investment performance? Standard economic theory posits that investors maximize the expected utility of their lifetime consumption stream by dynamically adjusting their portfolio allocations based on their wealth and expectations of the future. Although this way of modeling investors’ behavior may be close to reality for sophisticated investors such as mutual fund managers, it is questionable whether less sophisticated investors behave this way. For instance, some empirical evidence suggests that the majority of individual investors tend to buy stocks and then close off the entire position at some point in the future. Furthermore, contrary to the notions underlying theories like the APT or CAPM, individual investors seem to care more about an asset’s deviation from a reference point rather than rates of returns. When individual investors talk about their performance they tend to focus on individual assets rather than their portfolios. You are much more apt to hear a small investor say, I bought Microsoft at $40 and sold it at $80 or our house was worth half a million at the height of the market and now it is worth only $400,000. Not only are those not rates of return, but rates of return cannot even be determined without further information.

Earlier research indicates that individual investors think of their investing experience as a series of separate episodes in which they either made or lost money. That an asset was useful in providing diversification or hedging a risk or that the profits were realized over periods of different durations is less important. Furthermore, a growing amount of research shows that investors do not behave in the ways that expected utility theory predicts. In particular, the independence axiom seems troublesome as does the assumption of risk aversion, at least for losses. The latter assumption is not a requirement of expected utility theory, but it or something like it is important for most equilibrium models which follow from maximizing behavior.

Behavioral literature suggests alternative views of modeling investors’ behavior. Shefrin and Statman (1985) use “mental accounting” to justify investors’ concentrating on specific separate incidents. Thaler (1999) says that “A realized loss is more painful than a paper loss.” Kyle, Ou-Yang, and Xiong (2006) and Barberis and Xiong (2012) study models which assume the primary source of utility comes in a burst when a gain or loss is realized, and Frydman et al. (2011) find evidence using the neural data that supports the “realization utility” hypothesis.

In this paper, we use those notions to develop an intertemporal model of investors who have prospect theory’s S-shaped utility and who evaluate their own performance incident by

---

1 Feng and Seasholes (2005) show that among the 21,631 positions both opened and closed at a large brokerage house in the People’s Republic of China between January 1999 and December 2000, 69.15% of them consisted of a single purchase and a single sale — that is, investors seem to be investing in single stocks rather than rebalancing their portfolios.

2 More discussion on the psychological foundation of viewing investments as episodes can be found in Barberis and Xiong (2012).

3 Kyle, Ou-Yang, and Xiong (2006) study a one-time liquidation problem while Barberis and Xiong (2012) are concerned with continuation payoff and reinvestment.
incident based on realized profits and losses. Although it would be tempting to combine this model with a consumption-based model and solve the portfolio choice problem, we believe that studying realization utility alone is a good starting point for gaining a better understanding before moving toward a more general framework.

Our paper includes Barberis and Xiong (2012) as a special case, but unlike their model, our model generates voluntary sales both at gains and losses. This allows us to better match the empirical data to explain the disposition effect that people have higher propensity to sell stocks at gains than at losses. In addition, our model predicts that investors may be risk-seeking in some circumstances which helps explain a flatter security market line as shown in Black, Jensen and Scholes (1972), Fama and MacBeth (1973), and others. Our model also predicts the separate timing of trades at gains and losses and the prices at which these trades will be executed. However, our model typically predicts less frequent sales at losses than the empirical evidence suggests, and this indicates that realization utility alone may not fully explain individual investors’ trading patterns. Other motivations to sell or belief-based explanations⁴ need to be considered to complement our work.

The plan of our paper is as follows. In Section 2, we define incremental utility and discuss the distinction between it and expected utility theory. We also define the specific utility forms that are used throughout the paper. Section 3 analyzes the base case realization utility model. Sections 4 and 5 examine the value function and the optimal sales policies, respectively. In Section 6, we extend the model by considering taxes, a stochastic reference level, ongoing evaluation utility, and exogenously forced sales. Section 7 presents model predictions, and Section 8 gives some concluding remarks and direction for future research.

2. Reference-Dependent Incremental Utility in an Intertemporal Setting

A key distinction between Expected Utility Theory (EUT) and reference dependent behavioral theories (RDT) like Cumulative Prospect Theory (CPT) is that the former is based on the level of wealth (or consumption) while the latter is concerned with its deviation from a reference level. This distinction is not too important in a single-period setting where the reference level is fixed or otherwise exogenous so the dependence of utility on the reference level is imbedded in the utility itself and only comparative static analysis can be applied. However, in a multi-period setting, the effects of a changing reference level become important.

In a multi-period setting, EUT traditionally uses dynamic programming to determine the derived or indirect utility function. Then the effect of a change in wealth is measured by the change in this function. We could do something similar with RDT utility based on the gain or loss over the entire horizon; however, this seems at odds with the principle focus on changes. It seems more appropriate to define the utility function as some aggregation of all the individual changes.

One obvious procedure is to perform periodic evaluations with the reference level

⁴ Ben-David and Hirshleifer (2011) provide empirical evidence that can be linked to both preference-based and belief-based models.
updated to reflect the previous evaluation and then sum the utility contributions of the separate evaluations.\(^5\) In this context total expected lifetime utility is

\[
Y = \mathbb{E} \left[ \sum_{i} \Omega(\Delta X_i, R_i, t_i) \right] \quad \Delta X_i \equiv X_i - R_i
\]  

(1)

where \(X_i \equiv X(t_i)\) and \(R_i \equiv R(t_i)\) are the evaluated amount and reference level at the evaluation time \(t_i\). We impose \(\Omega(0, R, t) = 0\) as a centering condition. Depending on the scope of framing, \(X\) and \(R\) might be all of wealth or include only the value of those assets actively traded at each point.\(^5\) This needs to be coupled with a rule for updating the reference level. An obvious choice is the previous evaluation, \(R_{i+1} = X_i\), but this is not the only one. In addition, the evaluation times can be exogenous or endogenous. For example, realization utility might be assumed so utility is incremented only when a gain or loss is actively realized by a sale or some other action. We will refer to an intertemporal evaluation like (1) as \emph{reference-dependent incremental utility} or simply \emph{incremental utility}.

One empirical fact is that investors voluntarily sell at losses. First we would like to examine under what condition(s) our model will generate this result. The following theorem provides sufficient conditions for avoiding voluntary realization utility sales at losses.

\textbf{Proposition 1:} Assume that investors maximize lifetime expected utility as in (1), that the updating rule for the reference level is \(R_{i+1} = X_i\), and that there are no transactions costs. Then the following four conditions are jointly sufficient to preclude the voluntary realization of sales at losses

\[
\begin{align*}
(i) & \quad \frac{\partial \Omega}{\partial \Delta X} > 0 \quad (ii) & \quad \frac{\partial^2 \Omega}{\partial \Delta X^2} < 0 \quad (iii) & \quad \frac{\partial^2 \Omega}{\partial \Delta X \partial t} < 0 \quad (iv) & \quad \frac{\partial^2 \Omega}{\partial \Delta X \partial R} \geq 0.
\end{align*}
\]  

(2)

In addition, investors will voluntarily sell immediately to realize any gains.

\textbf{Proof:} See the appendix.

\hfill \blacksquare

A proof of this theorem is given in the Appendix. The intuition can be clarified by Figure 1 which illustrates a series of transactions that includes losses. It can be seen that each loss in the series can be divided into portions and matched against their counterparts among the gains. Each portion of every loss is matched against a portion from a later gain(s). The corresponding loss and gain portions are of equal size, but the gain portion has smaller marginal utility throughout due to risk aversion, its occurs later, and its lower reference level. Each of these properties makes the total utility from the gains less than the disutility of the losses, based on properties (ii), (iii), and (iv), respectively. This portion of the proposition remains valid even if there are transactions costs that are concave in the size of the transaction.

\(^5\) This is in keeping with EUT if the utility at each evaluation point is simply the change in the indirect utility function since the previous point. In this case the sum of the changes will equal the difference between the initial utility level and the utility of terminal wealth when there is no consumption.

\(^6\) In general each utility increment might also depend on the wealth. We suppress that dependence in this paper. If the investment holds all of wealth, then the reference level and the change are sufficient statistics for wealth.
To see that gains should be realized immediately, note that the marginal utility of gains is decreasing in magnitude, decreasing in time and increasing in its reference level so any gain has higher total utility if taken in two parts. In particular, in the absence of transaction cost, a gain of $X - R$ at time $t$ could have been taken as gains of $x - R$ at some earlier time $t_x$ and a gain of $X - x$ at time $t$ for total utility $\Omega(X-x, x, t) + \Omega(x-R, R, t_x) > \Omega(X-R, R, t)$. The reason transactions costs makes delaying the realization of gains optimal is that a large gain cannot be split into two parts with the same total gain. A transaction cost will reduce the investment and therefore the second part of the gain. Even if the costs are paid from an outside source, the disutility of paying them must be considered.

The first three conditions in (2) are standard properties. Presuming $X$ measures something desirable and the investor is risk averse, marginal utility is (i) positive and (ii) decreasing. The third condition is a generalization of positive time preference guaranteeing that the future is less important than the present. If utility has a separable discount factor, $\Omega(\Delta X, R, t) = f(t)v(\Delta X, R)$, then (iii) is equivalent to a decreasing $f(t)$. The effects of the reference level on utility are not so immediately intuitive. A positive $R$ derivative does not mean that the investor prefers investments with a larger reference level. What it indicates is that gains of a fixed size have a greater impact on incremental utility for investments with a higher reference level, while losses of a fixed size have a smaller impact; that is, $|v|$ increases with $R$ for gains and decreases for losses. Conversely, if $\partial^2 \Omega / \partial(\Delta X) \partial R$ is positive, then each extra dollar has greater impact for higher reference levels for both gains and losses.

It has been claimed that the disposition effect is a result of S-shaped utility function, concave for gains and convex for losses. The reasoning goes that, unlike for gains, the investor is willing gamble over the size of losses due to risk preference and therefore postpones realizing them. However, we see from Proposition 1 that this reasoning is exactly wrong. Losses are never taken with a concave realization utility function. It is only the convexity in the loss region that might make loss taking beneficial, and we show below that it can be part of an optimal policy.

Apparently for realization utility with reference-dependent preferences to generate both voluntary losses and delayed gains, we required transactions costs and marginal utility that is non-increasing in the reference level. Transactions costs are self-explanatory, but the effect of the reference level on utility calls for some further clarification.

It seems reasonable to assume that each dollar of a gain of a fixed size should typically have a smaller utility impact the larger is the reference level; that is, the marginal utility of gains should decrease with the reference level. This should also be true for losses if the investor is risk averse. For S-shaped utility, it is not as clear. For small losses the marginal utility probably should still decrease as the reference level rises; however, if the size of the loss is large relative to the reference level, this may no longer be true. An S-shaped utility function gets flatter the larger is the loss, and when the reference point is higher, there is a longer interval over which this flattening can occur before everything is lost, so the flattening might be gentler.

We refer to the property that the marginal utility gains decreases with the reference level as decreasing absolute preference scaling of gains (DAPSG). DAPSL is similarly defined for losses, and DAPS refers to both properties. If utility is differentiable with respect to $R$, then for
DAPS, $\frac{\partial^2 \Omega}{\partial (\Delta X) \partial R} < 0$. Like risk aversion, DAPS can be a local or a global property. As mentioned above we might not expect DAPSL to hold globally.

Models in finance are more commonly expressed in rate of return form so proportional preference scaling (PPS) might seem a more natural concept than absolute preference scaling. PPS is based on the gain or loss per dollar of the reference level and can be more conveniently represented by $\Omega(\Delta X, R, t) = e^{-\delta t} U(\Delta x, R)$ where $\Delta x \equiv \Delta X/R$. Utility has increasing, constant, or decreasing proportional preference scaling as $\frac{\partial^2 U}{\partial (\Delta x) \partial R}$ is positive, zero, or negative. For PPS either sign seems reasonable.

A further convenient, though restricting, assumption is that utility is multiplicatively separable in the scale and the proportional gain or loss, $U(\Delta x, R) = s(R) u(\Delta x)$. The function $u$ determines risk preferences over proportional gains and losses. It is increasing and either concave or, for CPT-like utility, increasing and S-shaped. Utility displays relative risk aversion that is independent of scaling and exhibits increasing, constant, or decreasing proportional preference scaling as $s$ is increasing, constant, or decreasing.

A simple scaling function is $s(R) = R^\beta$. For this form, the elasticity of utility with respect to the reference level is constant; that is

$$\frac{R \partial U}{U \partial R} = \beta .$$  \hfill (3)

For analytical convenience, this is the scaling assumption we adopt for this paper. The scaling parameter $\beta$ can be positive or negative giving increasing or decreasing PPS respectively.\(^7\) For DAPSG and DAPSL, we require

$$\frac{\partial^2 \Omega}{\partial (\Delta X) \partial R} = e^{-\delta t} R^{\beta - 2} u'(\Delta x) \left( \beta - 1 - \frac{\Delta x \cdot u^*(\Delta x)}{u'(\Delta x)} \right) < 0$$  \hfill (4)

for $\Delta x > 0$ and $\Delta x < 0$, respectively.

One convenient utility form that we will adopt for some of our analysis is scaled-Tversky-Kahnemann utility (scaled-TK)

$$s(R) = R^\beta \text{ } u_{\text{skT}}(\Delta x) = \begin{cases} (\Delta x)^{\alpha_G} & \Delta x \geq 0 \\ -\lambda (-\Delta x)^{\alpha_L} & \Delta x < 0 \end{cases}$$  \hfill (5)

with $0 < \alpha_G, \alpha_L \leq 1$ and $\lambda \geq 1$. In a single-period model, this is just a reinterpretation of Tversky-

\(^7\) In the next section we must impose $\beta \geq 0$ as a participation constraint for the model studied; however, this restriction need not necessarily hold in other contexts.
Kahnemann utility\(^8\) for our incremental utility model. As with CPT, the parameters \(\alpha_G\) and \(\alpha_L\) determine the investor’s risk aversion over gains and risk preference over losses. Loss aversion is measured by \(\lambda\); if \(\alpha_G = \alpha_L\), then the disutility of a loss will be \(\lambda\) times as large as the utility of a gain of the same size. This utility function displays DAPSG (DAPSL) if \(\beta < \alpha_G\) (\(\beta < \alpha_L\)).

Two problems with scaled-TK utility are that the risk aversion over gains is quite limited by the restriction \(\alpha_G > 0\) and that marginal utility is very high for small gains or losses — indeed marginal utility is infinite at 0 (for any \(\alpha_G < 1\) and \(\alpha_L < 1\)). These properties make it difficult to calibrate TK utility to many of the risk-return trade-offs observed in the market. Therefore, we also consider the following utility specification, called modified-TK utility, which does permit calibration to larger risk aversions.

\[
s(R) = R^\beta 
\]

\[
u_{mKT}(\Delta x) = \begin{cases} 
[(1 + \Delta x)^{\alpha_G} - 1]/\alpha_G & \Delta x \geq 0 \\
-\lambda[1 - (1 + \Delta x)^{\alpha_L}]/\alpha_L & -1 \leq \Delta x < 0. 
\end{cases}
\] (6)

For \(\alpha_G < 1 < \alpha_L\), the utility will be S-shaped like TK-utility.\(^9\) As \(\alpha_G\) can take any value less than one, high risk aversion over gains is possible. Marginal utility is bounded at \(\Delta x = 0\). It reaches the value \(\lambda\) just below zero and 1 just above zero displaying a true kink there with a discontinuous change in marginal utility. Modified-TK utility exhibits DAPSG if \(\beta < \alpha_G\).\(^10\)

Once the utility function, \(u\), is known, the scaling parameter, \(\beta\), can be determined by answering a question like the following: “Gaining $10 relative to a reference level of $100 makes just as happy as gaining \(z\) relative to a reference level of $1000.” Given the answer to this question, \(\beta = \log[u(0.1)/u(z/1000)]\). If losses scale differently from gains, a similar question would be required about losses. Note that the loss aversion parameter, \(\lambda\), is not involved in this determination. Figure 2 illustrates this procedure plotting \(\beta\) against \(z\) for scaled-TK utility with various values of \(\alpha\) and modified-TK utility with \(\alpha = -1\).\(^11\)

There is another advantage of this specification. For \(\beta = 0, \lambda = 1,\) and \(1 - \alpha_G = \alpha_L - 1,\) (i.e., risk aversion of gains equal to risk preference of losses), a gain followed by a loss that returns the investor’s wealth to the original level leaves him with zero net utility (apart from

---

8 This utility function was introduced in Tversky and Kahneman (1992). For \(\beta = \alpha_G = \alpha_L,\) utility of a single gain or loss is \((X - R)^\alpha\) or \(-\lambda(R - X)^\alpha\). Using different scaling functions for gains and losses with exponents \(\beta_L = \alpha_L\) and \(\beta_G = \alpha_G,\) results in TK utility with distinct exponents for gains and losses. The Barberis and Xiong (2012) model is the special case \(\beta = \alpha_G = \alpha_L = 1.\)

9 As usual \(\alpha_G = 0\) corresponds to \(u(\Delta x) = \ln(1 + \Delta x).\) Modified-TK utility can also be adapted to study strictly risk-averse incremental utility by setting \(\alpha_L < 1.\) Utility is obviously increasing and strictly concave. (Marginally utility at \(\Delta x = 0\) is 1 and at 0 is \(\lambda\) so utility is strictly concave if \(\lambda > 1.\)) If \(\alpha_L = \alpha_G\) and \(\lambda = 1,\) this is incremental power utility; otherwise there is a discontinuous change in risk aversion (if \(\alpha_G \neq \alpha_L\)) or in slope (if \(\lambda \neq 1\)) at 0.

10 Assuming \(\beta < \alpha_G < 1 < \alpha_L,\) modified-TK utility exhibits DASPL for small losses and IASPL for large losses with \(\Delta x < -(1 - \beta)(/\alpha_L - \beta).\)

11 The answer \(z\) to the given question is almost invariant to risk aversion for modified-TK utility since \(\beta = \log[u(1)/u(z/1000)] \approx 2 - \log z.\) So, for example, an answer of $50 would imply \(\beta = 0.296\) for \(\alpha = 0.5\) and \(\beta = 0.262\) for \(\alpha = -3\) while the approximation gives \(\beta \approx 0.301.\)
discounting). This mimics the property for offsetting dollar changes in wealth of TK utility when its risk aversion and risk tolerance are equal and \( \lambda = 1 \).

In addition to changing upon evaluation, the reference level might also change between evaluations.\(^{12}\) For example, the reference level might reflect lost opportunities and rise over time at the interest rate, the rate of inflation, or a subjective rate that the investor believes he should earn on the investment. In this context the reference point is serving as a goal rather than an anchor. We can model all of these possibilities by assuming that the reference grows at the exogenous constant rate \( \zeta \). As stated in Barberis and Xiong (2012), modeling investors’ behavior using a realization utility framework is related to less sophisticated individual investors, and this group of investors is more likely to follow a simple rule like a constant growth to update their reference point over time. Another possibility is that the reference level grows randomly reflecting the actual return on an alternative investment that was foregone. We will also examine this case later.

3. The Realization Utility Model

This section determines and evaluates the investment strategies of an investor who has an incremental utility function as described above and invests in one of a set of stocks with identically distributed returns. Though we refer to the underlying assets as stocks, they could be portfolios or any kinds of project. The investor takes a position in one of the stocks holding it until he decides to sell. Upon sale he realizes an incremental utility burst based on the difference between the sales price and the reference level. This is the primary source of utility though in some specifications considered below other factors may also contribute. After a sale, the investor reinvests the entire net proceeds in another stock with an identical distribution.\(^{13}\) The investment may, but need not, be the investor’s entire wealth. If it is not his entire wealth, we assume the investor narrowly frames his utility assessment to evaluate the utility from this investment in isolation from others.

The underlying stock price evolves as \( dS/S = \mu dt + \sigma d\omega \). The parameters, \( \mu \) and \( \sigma^2 \), are the expected instantaneous growth rate and proportional variance. The investor holds some number of shares whose total value, \( X \), remains proportional to the stock price, \( S \), between sales. His value function, \( \Phi(X, R, t) \), is the expected discounted value of all future utility bursts from sales. Since \( \Phi \) is an expectation, by the law of iterated expectations and Ito’s lemma

\(^{12}\) When the reference rate changes over time in the absence of a trade, there are two interpretations to gains and losses. The investor will feel a gain (loss) whenever he sells the asset at a price exceeding (less than) the reference level. However, an observer, like an econometrician, is generally not privy to the investor’s feelings and presumably records gains and losses relative to the initial purchase price. Throughout this paper we adopt the first interpretation. For empirical work, the latter would obviously be required.

\(^{13}\) This setup is a simplified one to concentrate on realization utility. The assumptions of holding a single stock and of full reinvestment allow us to use a perpetual model and eliminate the need to consider diversification and other portfolio considerations.
\begin{equation}
0 = \mathbb{E}\left[ dY(X, R, t) \right] = \mu X Y_X + \frac{1}{2} \sigma^2 X^2 Y_{XX} + \zeta R Y_R + Y_t
\end{equation}

(7)

for periods during which there are no sales.

On occasion the investor sells his stock and realizes a gain or loss with a corresponding positive or negative utility burst. We assume there are proportional transactions costs of \( k_s \) and \( k_p \) on sales and purchases, respectively.\(^{14}\) Upon a sale, the investor receives a utility burst based on the subjective realized gain or loss. We denote the subjective gain as \( \Delta X \equiv \kappa X - R \) or the proportional gain as \( \Delta x \equiv \kappa X/R - 1 \) where \( \kappa \) is a parameter indicating how transactions costs affect the investor’s perception of gains and losses. The utility burst at a sale is

\[ U(\Delta x, R) = R^0 u(\kappa X/R - 1). \tag{8} \]

After the investor sells, reinvests, and pays both transaction costs, he has \( [(1-k_s)/(1+k_p)]X \equiv KX \) invested.

There are several ways that the investor might view his gain or loss subjectively. The first is that he fully recognizes transactions costs and compares the net reinvested amount with the reference level for a gain of \( KX_1 - R_0 \). A second is that he views the gain as the difference between the net proceeds of the sale and the reference level, \( (1-k_s)X_1 - R_0 \). Finally, he might ignore transactions costs completely and view his gain as \( X_1 - R_0 \). These three cases are covered by setting the parameter \( \kappa \) to \( K \), \( 1-k_s \), or \( 1 \), respectively.\(^{15}\) For example, suppose the investor sells for \$130 stock for which he has a reference level of \$100. If both transactions costs are 1%, he will receive \$128.70 and be able to reinvest \$127.43. For the three subjective views discussed above, he would realize a gain of \$27.43, \$28.70, or \$30. Intermediate views are also possible, particularly if the transactions costs have multiple parts such as a bid-ask spread and a commission. We leave the parameter \( \kappa \) free allowing many interpretations.

There is also ambiguity as to how the investor sets the new reference level. He might take the reference level to be the net amount invested, \( KX_1 \), or he might view it as the gross amount invested including the purchasing cost, \( (1-k_s)X_1 \); that is, \$127.43 or \$128.70 in the example above. Again it might be some intermediate amount. In our analysis, we will assume the investor fully accounts for costs and sets the reference level to the net amount invested; \( R_1 = KX_1 \).\(^{16}\)

---

\(^{14}\) When an investor buys a share of stock for \( S \), he pays a total of \( S(1+k_p) \), and when he sells a share for \( S \) he receives \( S(1-k_s) \). Unlike in most models, we sometimes need to distinguish these costs depending on how the investor assesses his utility and cannot simply state a round-trip transactions cost. Due to the transactions costs, the interpretation of \( X \) must now be changed slightly. \( X \) represents the value of the assets held for investment, but \( \Delta X \) is no longer the subjective gain or loss.

\(^{15}\) Barberis and Xiong (2012) adopt the first convention that \( \kappa = K \); however, we will leave the parameter \( \kappa \) free. We believe that most investors do not consider the reinvestment cost when viewing their gain so that \( \kappa = 1-k_s \) or even \( \kappa = 1 \) is more likely.

\(^{16}\) Barberis and Xiong (2012) also adopt this interpretation which in their notation is \( R_1 = (1-k)X_1 \) with \( k \) being the round-trip transactions cost. If an investor adopts the gross cost view for his reference level, then equation (9) has \( Y(XX, (1-k_s)X, t) \) as the final term on the right-hand side. The final terms on the right-hand side of (12) are
Equating the value function before the sale to the utility burst of the sale plus the post-reinvestment value function

$$\hat{Y}(X, R, t) = e^{-\delta t}U(\Delta x, R) + \hat{Y}(KX, KX, t).$$

(9)

The optimal sales strategy can be characterized as sell for a gain or loss at $X(R, t)$ and $X(R, t)$, respectively. The strategies cannot depend on the current value of $X$. The investor cannot, for example, plan to sell at $X < \hat{X}$ when $X < \hat{X}$, only to change his mind and decide sell at $X > \hat{X}$ when he reaches $X = \hat{X}$. There is no mechanism in our model to permit such time inconsistency. However, the lower optimal sales points need not exist. Under some circumstances selling at a loss is never optimal. This issue is examined further below. An upper optimal sales point must exist, at least in the limit, since some sales at gains must be made to create a positive $\hat{Y}$ otherwise the investor will not enter the market.

The utility bursts in (8) are homogeneous of degree $\beta$ in $X$ and $R$ for either specification. Since the reference level has a constant growth rate, and the stock price process has stochastic constant returns to scale, the value function, $\hat{Y}$, must also be homogeneous of degree $\beta$ in its arguments. Furthermore, for our perpetual problem, the future looks the same depending only on $R$ and $X$ so time affects the value function only through discounting, and we can define a time-independent value function $v(x) = e^{\delta t}R^\beta \hat{Y}(X, R, t)$ where $x = X/R$. This homogeneity property requires that $\beta$ be nonnegative. Since $\hat{Y}(X_0, X_0, t) = X_0^\beta \hat{Y}(1,1,t)$, the initial utility is decreasing in the original investment amount, and the investor would never participate — always preferring to invest less than he actually had.

In terms of the variable, $x$, the reduced form of the value function, $v$, is the solution to the reduced form of the relation in (7), an ordinary differential equation

$$0 = \frac{1}{2} \sigma^2 x^2 v'' + (\mu - \zeta) xv' - (\delta - \zeta \beta)v.$$

(10)

This equation resembles the Black-Scholes option equation and some insights can be derived from that even without a solution. A growing reference level ($\zeta > 0$) has an effect similar to decreasing the expected growth rate since it is less likely that the investment will outperform the growing reference level. But $\zeta$ also affects the discounting. Since each utility burst is weighted by $R^\beta$ and the reference level grows at the rate $\zeta$, the effective discount rate is reduced by the product $\zeta \beta$.

[[1-k_2] \Theta v(K(1-k_2))] and [(1-k_2) \Theta v(K(1-k_2))]. Most brokerage accounts show the purchase price which tends to emphasize the net after-cost value as the reference level. On the other hand, the taxable capital gain excludes the purchase cost which tends to emphasize the gross cost.

17 Unique optimal plans do not exist for some parameter values. If utility is unbounded above and the expected growth rate on the stock is high compared to the subjective discount rate, there can be a transversality violation since waiting forever to sell leads to infinite utility just as in the standard investment problem. There also can be a second transversality violation in this problem when $\beta$ is negative or large. In either case there are multiple plans that lead to infinite utility. The transversality conditions, (i) $\beta \zeta < \delta + \rho$, (ii) $\beta < \gamma_1$, (iii) $\alpha_\zeta < 0$ or $\alpha_\zeta [\mu - \zeta + (\alpha-1)\sigma^2/2] < \delta + \rho - \beta \zeta$, are all derived in the Appendix. The parameter $\rho$ is introduced in some extensions to the model; for now $\rho = 0$. 


The general solution to (10) is

\[ \nu(x) = C_1 x^{\gamma_1} + C_2 x^{\gamma_2}, \quad \text{where} \quad \gamma_{1,2} = \frac{\zeta - \mu + \frac{1}{2} \sigma^2 \pm \sqrt{(\mu - \zeta - \frac{1}{2} \sigma^2)^2 + 2(\delta - \zeta \beta)\sigma^2}}{\sigma^2}. \] (11)

Again due to the homogeneity, the optimal sale strategy must be to realize a gain or loss when the stock price reaches a constant multiple or fraction of the reference level; \( X(R) = \Theta R \) and \( X(\overline{R}) = \theta \overline{R} \) or \( \overline{x} = \Theta \) and \( \overline{x} = \theta \). The upper sale point, \( \Theta \), must exceed \( \frac{1}{\kappa} > 1 \) as otherwise the sale is not a gain. The lower sale point, \( \theta \), must be less than 1. While a sale at a point in the range \( (1, \frac{1}{\kappa}) \) would produce a subjective loss after the transactions cost, such a sale could never be part of an optimal strategy. If the investment were always sold at a price in this range, there would never be any sales at a higher price as the stochastic process for \( x \) is continuous and begins at 1. But this means that only losses with negative utility bursts would be realized so \( \nu \) which is their present value would be negative and could not be the optimized value.

Upon each type of sale, the boundary conditions from (9) are

\[ \nu(\Theta) = R^{-\beta} U(\kappa \Theta - 1, R) + e^{\delta t} R^{-\beta} \gamma(K \Theta R, K \Theta R, t) = u(k \Theta - 1) + (K \Theta)^\beta \nu(1) \]
\[ \nu(\theta) = R^{-\beta} U(\kappa \theta - 1, R) + e^{\delta t} R^{-\beta} \gamma(K \theta R, K \theta R, t) = u(k \theta - 1) + (K \theta)^\beta \nu(1). \] (12)

Substituting from (11), the constants, \( C_1 \) and \( C_2 \) satisfy

\[ C_1 \gamma_1 + C_2 \gamma_2 = u(k \Theta - 1) + (K \Theta)^\beta (C_1 + C_2) \]
\[ C_1 \gamma_1 + C_2 \gamma_2 = u(k \theta - 1) + (K \theta)^\beta (C_1 + C_2) \] (13)

whose solution is

\[ C_1 = \frac{c_2(\Theta)u(k \Theta - 1) - c_2(\theta)u(k \theta - 1)}{c_1(\Theta)c_2(\Theta) - c_1(\theta)c_2(\theta)} \quad \text{and} \quad C_2 = \frac{c_1(\Theta)u(k \theta - 1) - c_1(\theta)u(k \Theta - 1)}{c_1(\Theta)c_2(\Theta) - c_1(\theta)c_2(\theta)} \quad \text{where} \quad c_1(\phi) = \phi^{\gamma_1} - (K \phi)^\beta. \] (14)

The optimal sales points, \( \Theta \) and \( \theta \), can be determined by maximizing \( C_1 + C_2 \) which is the value of \( \nu(1) \), or by applying the smooth-pasting condition at both points.\(^{19}\) Note that the smooth pasting condition is not simply \( \nu' = u' \), but must be applied to the reduced equivalent of (9) which has the continuation value as well as the utility burst from the sale on the right hand side. At the two boundaries

\(^{18}\) See the Appendix for more details on the constancy of the optimal policy.

\(^{19}\) The optimal sales strategy must maximize the value function at every value of its argument in the continuation region. \( X = R \) or \( x = 1 \) is guaranteed to be in the continuation region since \( \theta < 1 \) and \( \Theta > 1/\kappa > 1 \). Some care must be taken in applying the smooth-pasting condition. In some cases there are multiple local maxima and in other cases, there are corner optima as illustrated in Figure 3 below.
\[
\gamma_1 C_2 \Theta^{-1} + \gamma_2 C_2 \Theta^{-1} = \kappa u'(\kappa \Theta - 1) + \beta K^\beta \Theta^{\beta-1} (C_1 + C_2) \\
\gamma_1 C_1 \Theta^{-1} + \gamma_2 C_2 \Theta^{-2} = \kappa u'(\kappa \Theta - 1) + \beta K^\beta \Theta^{\beta-1} (C_1 + C_2).
\]  
(15)

The solution and the optimal \( \Theta - \Theta \) strategy is illustrated in Figure 3. The no-sales region runs from \( \theta \) to \( \Theta \). The value function \( v(\Theta) \) is tangent to the payoff \( u(\kappa \Theta - 1) + (\kappa \Theta)^\beta v(1) \) at \( x = \Theta \). A similar relation holds at the loss point \( \theta \).

One obvious question is why an investor should ever realize a loss and take a negative utility burst rather than just wait as long as needed for the stock price to rise above the reference point and only take positive utility bursts. In fact with piecewise linear utility, he would never do so, and Barberis and Xiong (2012) introduced random liquidity shocks to force some sales at losses. However, with our S-shaped utility specification, a voluntary sale at a loss can be optimal. For example, with scaled-TK utility parameters, \( \alpha_G = \alpha_L = 0.5, \lambda = 2, \beta = 0.25, \delta = 4\%, \zeta = 5\% \) and economic parameters \( \mu = 10\%, \sigma = 30\%, k_s = k_p = 1\%, \kappa = 1-k_s, \) the optimal strategy is \( \Theta = 1.033, \theta = 0.417 \) with a value of \( v(1) = 8.40 \). The optimal one-point policy is \( \Theta = 1.031 \) with a value of 7.25.

The reason for the superiority of the two-point strategy is the benefit of resetting the reference level near the current stock price upon a sale. This resetting accomplishes two things. It makes a sale at a gain more likely in the near future (which is also true under linear utility) and also increases the expected utility of a sale since the marginal utility of \( u \) is highest at \( \Delta x = 0 \). Of course future negative utility bursts would also be larger, but  the investor can postpone those until there is a larger loss with a lower marginal utility. That is, with S-shaped utility, the investor is willing to trade one large-dollar loss for several small-dollar gains at high marginal, and therefore total, utility. This benefit is larger the steeper is the marginal utility of small gains so sales at losses are more important for scaled-TK utility with \( u(0+) = \infty \). At the opposite extreme, with piecewise linear utility where the marginal utility for any gain is the same and smaller than that for losses so there are no voluntary sales at losses.

Of course, in some cases, it may never be optimal to realize a loss. For example, if we change \( \beta \) to 0.5 in the previous example, then the optimal sales points are \( \theta = 0, \Theta = 1.374^{20} \). Whether a one-point or two-point sales policy is optimal depends on the various utility and economic parameters. However, this problem is not a standard convex optimization. As shown in Figure 4, the reduced value function after a sale, \( v(1) \), is not a concave function of the lower sales point \( \theta \). Both an interior local maximum and a corner local maximum at zero are possible. This shape of the value function is due to three factors: i) the high marginal disutility of small losses, ii) the option value of resetting the reference point for future gains, iii) the scaling factor that

---

20 If we assume, ex ante, that there are only sales at gains, then \( C_2 = 0 \) to keep \( v \) bounded since \( \gamma_2 < 0 \). In this case only the first equation in (13) applies and \( v(0) = C_1 0^{\gamma_1} = 0 \). Note that for other stochastic processes in which zero is an accessible absorbing boundary, two alternative boundary conditions can be applied at 0. Since the price must remain at 0, no future gains are possible. So \( v(0) = u(-1) \) if the investors “sells” or otherwise feels he’s taken this loss. If the investor continues to hold the worthless asset and does not feel the loss, then \( v(0) = 0 \). For the lognormal diffusion used here, 0 is a singular point of the differential equation and no boundary condition apart from boundedness can be imposed.
reduces continuation utility after a sale at a loss. The first factor means that one large loss is better than several smaller losses so the value function is decreasing in $\theta$ for $\theta$ near 1. However, the second factor makes realizing *some* losses beneficial; this increases the value function in the mid-range of $\theta$. The third factor makes avoiding any loss better than taking a large loss so the value function is again decreasing in $\theta$ for $\theta$ near zero.

Figure 4 illustrates three cases, plotting $v(1)$ for three values of $\lambda$. The upper sales point is fixed at its distinct optimal value in each case. For $\lambda = 2.8$, it is optimal to sell for a loss at $\theta = 0.19$. For $\lambda = 3$, there is an apparent optimum at $\theta = 0.12$, but this is only a local maximum as never selling at a loss provides higher utility as shown. For the critical value of $\lambda \approx 2.892$, both selling for a loss at $\theta = 0.16$ and never selling for a loss provide the same utility. Note that the lower sales point, $\theta$, does not decrease smoothly to zero as $\lambda$ is lowered; $\theta$ drops discontinuously from 0.16 to 0. A similar change in regime is true for the other parameters. The two regimes are characterized in Proposition 2.

**Proposition 2:** Scaled-TK utility has both an upper and a (non-zero) lower optimal sales points if and only if $\lambda$ is less than the critical value

\[
\lambda_* = \frac{(\kappa \Theta_* - 1)^{\alpha_* - 1}}{(1 - \kappa \Theta_*)^{\alpha_* - 1}} \left( \frac{\Theta_*}{\Theta_*} \right)^\beta (\alpha_G - \gamma_1) \kappa \Theta_* + \gamma_1
\]

where $\theta_*$ and $\Theta_*$ are the solutions to the non-linear equations

\[
0 = (\alpha_G - \gamma_1) \kappa \Theta_*^{\gamma_1 - 1 + \beta} + \gamma_1 \Theta_*^{\gamma_1 - \beta} - (\alpha_G - \beta) K^\beta \kappa \Theta_* - \beta K^\beta
\]

\[
0 = (\alpha_L - \gamma_1) \kappa \Theta_*^{\gamma_1 - 1 + \beta} + \gamma_1 \Theta_*^{\gamma_1 - \beta} - (\alpha_L - \beta) K^\beta \kappa \Theta_* - \beta K^\beta.
\]

The reduced-form valuation function is given in (11) where $C_1$, $C_2$, $\Theta_*$ and $\theta_*$ are jointly determined by four equations in (13) and (15).

If $\lambda$ is greater than this critical value, only gains are realized. The reduced-form value function is $v(x) = C_1 x^{\gamma_1}$ for $x \in (0, \Theta)$ where $C_1$ and $\Theta_*$ are jointly determined by

\[
C_1 \Theta^{\gamma_1} = \mu (\kappa \Theta - 1) + (K \Theta)^\beta C_1
\]

\[
\gamma_1 C_1 \Theta^{\gamma_1 - 1} = \mu' (\kappa \Theta - 1) + \beta K^\beta \theta^{\beta - 1} C_1
\]

Modified-TK utility has both upper and lower optimal sales points if and only if $\lambda$ is less than the critical value

\[
\lambda_* = \frac{\alpha_L}{\alpha_G} \left( \frac{\theta_*}{\Theta_*} \right)^\beta (\alpha_G - \gamma_1) \kappa^{\alpha_G} \Theta^{\alpha_G} + \gamma_1
\]

\[
(\alpha_L - \gamma_1) \kappa^{\alpha_L} \theta^{\alpha_L} + \gamma_1
\]

where $\theta_*$ and $\Theta_*$ are the solutions to the non-linear equations.
\[ 0 = (\alpha - \gamma_l)k^{\alpha_L}z^{\alpha_L} + \gamma_l(z_0 - \beta)k^{\beta}(k\theta)^{\alpha_G} - \beta K^p. \]

The reduced-form valuation function, constants, and optimal sales points are again given by (11), (13), and (15). If \( \lambda \) exceeds this critical value, only gains are realized and the constant and optimal sales point, \( \Theta \), are determined by (18).

**Proof:** See the appendix.

The Barberis-Xiong (2012) model is a special case of either modified-TK or scaled-TK utility with \( \alpha_L = \alpha_G = \beta = 1 \). For this model, or indeed any realization-utility model with linear utility for gains and losses and \( \beta > 0 \), the critical value \( \lambda^* \) is less than 1. Therefore, the sales-at-gains-only policy dominates the two-point sales policy.

### 4. The Value Function

Figure 5 presents the reduced value function measured at the time of any reinvestment (i.e., \( v(1) \) evaluated at the maximizing values of \( \theta \) and \( \Theta \)) plotted against the asset’s expected growth rate, \( \mu \), and standard deviation, \( \sigma \). The value functions are normalized to 1 at the default values \( \mu = 10\% \) or \( \sigma = 30\% \).

The smaller is \( \beta \), the larger is the value of this option, and therefore \( v \), because after a sale at a loss, the reference level is reset lower by a factor of \( (K0)^{\beta} \), and the closer this reset is to 1, the higher is the value. Similarly, a larger \( \zeta \) keeps this option value larger because the asset value stays closer to the reference level when \( \zeta \approx \mu \) preserving the optionality. For sufficiently large \( \beta \), small \( \zeta \), or small \( \delta \), the risk aversion over gains predominates making \( v \) a

---

21 Standard utility functions are defined only up to a positive affine transformation. Realization utility has its level set so that no gains and losses gives a utility of zero, but scaling is still arbitrary.
decreasing function of $\sigma$. The option effect is also smaller for modified-TK utility than for scaled-TK utility. The option is worth less under modified-TK utility because it is “exercised” less frequently as the no-trade region is wider as discussed in the next section of this paper.

Because the value function is not necessarily decreasing in $\sigma$, the indifference curves for this utility may not have their normal shape which is increasing and convex for the standard mean-variance problem. Figure 6 plots the indifference curves for scaled- and modified-TK utility. For $\beta = 0$, the indifference curves are concave for large $\sigma$ and decreasing in some cases. For $\beta = 0.25$, the indifference curves are mostly increasing and convex but their slopes are shallower than typical Sharpe ratios.\(^{22}\)

5. The Optimal Sales Policies

Figures 7 and 8 present graphs of the optimal selling points for gains and losses for both scaled-TK and modified-TK utilities and for investors who set their reference level at the net or gross amount invested. For both utility functions and both reference settings, realized losses are typically larger than realized gains so the basic strategy is to realize a few large losses and many small gains.

In all cases, the no-sales region is wider for an investor who views a gain as the excess over the reference level of the net amount reinvested rather than the net amount realized on the sale; i.e., $\kappa = K = (1-k_s)/(1+k_p)$ rather than $\kappa = 1-k_s$. For the default parameters, the no-sale region for scaled-TK utility widens from $(0.417, 1.033)$ to $(0.292, 1.052)$. For modified-TK it widens from $(0.246, 1.110)$ to $(0.208, 1.141)$. While both investors face the same transactions costs, an investor who internalizes them more when assessing his well-being is obviously more reluctant to trade.

The no-sales regions are generally quite a bit narrower for scaled-TK utility than for modified-TK utility. For scaled-TK, the marginal utility of small gains and losses is very large, but both drop quickly as the magnitude of gain or loss increases. This is not true for modified-TK utility; marginal utility is never greater than 1 and marginal disutility is never greater than $\lambda$. This means that a series of small gains is much more valuable than a single gain of the same size for scaled-TK utility but only somewhat more valuable for modified-TK utility. So an investor with scaled-TK utility is more willing to ignore transactions costs and trade quickly at a gain. The same marginal utility comparison is true for losses, but the reason for making any trade at a loss is to set up future gains, and again scaled-TK utility is favored in this. For scaled-TK utility, but not modified-TK utility, the disutility of any loss can always be offset by some series of gains totaling the same size.\(^{23}\) For example, using the scaled-TK parameters for Figure 7 ($\alpha_L = \alpha_G = 0.5$ and $\lambda = 2$), the disutility of a 2% loss can be offset by taking a 2% gain in four 0.5% increases. But with the modified-TK utility in Figure 8 ($\alpha_L = 5$, $\alpha_G = -3$, $\lambda = 2$), there is no way to divide a 2% gain into small pieces to offset the disutility of a 2% loss.

---

\(^{22}\) The maximum slope illustrated is about 0.2 which is about half of the market’s annual Sharpe ratio.

\(^{23}\) This is true even if $\alpha_L \neq \alpha_G$, provided only $\alpha_G < 1$. The example ignores transactions costs, but even with transactions costs, it is still possible to find a series of gains that offset a given loss.
This marginal utility effect makes it possible to offset the disutility of losses with gains, but the optimal way to do so depends on all the parameters including, in particular, $\alpha_L$. For scaled-TK utility with $\alpha_L$ near zero, marginal disutility remains large longer as the loss gets larger, but large losses have about the same disutility for any $\alpha_L$. In fact losing everything ($\Delta x = -1$) has a utility of $-\lambda$ regardless of $\alpha_L$. These factors make the investor wait longer to realize a loss as his risk-seeking behavior increases ($\alpha_L$ decreases from 1 to 0). With modified-TK utility, the utility is higher and the marginal utility is lower for any size loss the larger is $\alpha_L$ so loss-taking to set up future gains occurs more frequently with higher $\alpha_L$. We also see that there is a change in regime with $\theta$ dropping to zero for $\alpha_L \approx 0.4$. This is just another manifestation of the effect discussed in the previous section. This change in regime can also be seen in many of the other graphs. The optimal gain-taking policy is essentially unaffected by $\alpha_L$ though what little effect is there does also serve to widen the no-trade region for both utilities very slightly. This is a secondary effect of the change in $\alpha_L$.

There is also a distinct difference between the effects of $\alpha_G$ for the two utility functions. The upper sales point rises for both of them. For modified-TK utility, $\Theta$ rises slowly and then asymptotically as $\alpha_G \to \gamma_1$ where the transversality violation comes into play. For scaled-TK utility, $\Theta$ rises very slowly for its entire range. There is also a transversality violation for scaled-TK utility though there is no asymptotic rise. There is a significant impact on the optimal loss sales strategy for scaled-TK utility. As $\alpha_G$ approaches 0, the marginal utility of small gains increases rapidly making small losses “affordable” and desirable to set up future gains. With modified-TK utility, the marginal utility of small gains is never above 1, but any given gain does have somewhat higher utility the larger is $\alpha_G$, so losses become slightly more advantageous and $\theta$ rises slowly.

For both utility functions, the loss aversion parameter $\lambda$ has a significant effect on the optimal loss-taking strategy and a very minor effect on the optimal gain strategy. The more losses hurt (larger $\lambda$), the longer the investor waits before realizing either a loss or gain. It is obvious to wait longer before taking a harmful loss. The reason for waiting for a slightly higher price before taking a gain is that doing so resets the reference level and makes the more harmful losses arrive later on average.

The no-sales region widens as the scaling parameter, $\beta$, is increased. That the lower sales point, $\theta$, should decrease with $\beta$ is apparent. The continuation utility after a sale at a loss and a reset of the reservation level is $(K\theta)^{\beta} v(1)$ so the larger is $\beta$ the lower is this utility. This increases the investor’s reluctance to experience a loss just to make more future gain possible. It might seem that opposite would be true with a sale at a gain so increasing $\beta$ would encourage sales even more thereby decreasing $\Theta$. But that is not the only effect. The tradeoff can be seen most easily by ignoring sales at losses.

Consider an investor who follows a fixed $\Theta$ strategy. Assuming for simplicity that $\zeta = 0$, he sells for the first time when $X_t$ hits $\Theta R_0$, and after repurchasing, his reference level will be

$$24 \text{ The utility burst of a small gain is } \alpha_G^{-1}[(1 + \Delta x)^{\alpha_G} - 1] \approx \Delta x - \frac{1}{2}(1 - \alpha_G)(\Delta x)^2.$$
$K\Theta R_0$. Continuing in the same way, his reference level for the $n+1$st sale will be $(K\Theta)^n R_0$. His total lifetime expected utility is

$$
\mathbb{E} \sum_{n=1}^{\infty} e^{-\delta \tau} R_n^{\beta} u(\Delta x) = u(\Delta x) \sum_{n=1}^{\infty} R_n^{\beta} \mathbb{E}[e^{-\delta \tau}] = u(\Delta x(\Theta)) \frac{R_0}{(K\Theta)^{\beta}} \sum_{n=1}^{\infty} (K\Theta)^{\beta} \mathbb{E}[e^{-\delta \tau}].
$$

(21)

A larger $\Theta$ increases the time to each sale, which makes each expectation in the sum smaller due to more discounting, but if $\beta > 0$ there is an offsetting increase in value due to the $(K\Theta)^\beta$ term. As can be seen in the figures, this second factor dominates and $\Theta$ increases with $\beta$. Figure 7 illustrates that if $\beta$ is sufficiently large, the optimal sales point, $\Theta$, can become unboundedly large; the optimal $\Theta$ hits an asymptote as $\beta$ nears $\gamma_1 = 0.5124$. In fact, as shown in the Appendix, $\gamma_1$ is an upper bound on $\beta$ to keep lifetime utility finite for any utility function.

The effect of the subjective discount rate, $\delta$, on the optimal sales policy is unusual. A more impatient investor would want to realize gains sooner and defer losses longer. We see that $\Theta$ is decreasing in $\delta$ as expected; however, $\theta$ is not. The desire to take gains early induces a derived willingness to realize losses to set up these future gains. This causes $\theta$ to also be increasing in $\delta$ at low discount rates; however, at higher discount rates the initial intuition will dominate because the benefits of resetting occur farther into the future. The option value is greater for scaled-TK than modified-TK utility so the increasing portion of the $\theta$ curve is steeper and longer.

Decreasing the transactions costs, $k_s$ and $k_p$, narrows the no-sales region. Costs obviously reduce the advantage of selling either to realize a gain or to reset the reference level so trading increases as the costs are lowered. For scaled-TK utility, the trading frequency increases without limit as the costs go to zero. Marginal utility is infinite at $\Delta x = 0$, so the investor takes every opportunity to realize even the smallest of gains. Of course, there is unbounded marginal disutility for near-zero losses, but under the optimal strategy $\theta$ approaches 1 slower than does $\Theta$, so there is a net increase in the value function which becomes unbounded in the limit as shown in the Appendix. For modified-TK utility, marginal utility remains bounded at zero so the value function cannot be made infinite even in the limit of zero transactions costs. Nevertheless, marginal utility is at its highest at $\Delta x = 0$, so it is always better to take several small gains rather than a single larger gain when the extra cost of doing so is negligible, and in the limit as the costs go to zero, gains are realized immediately. However, the frequency for realizing losses does remain bounded, and the optimal $\theta$ has a limit near 0.45. Since the marginal utility of a gain is never larger than one, the investor cannot offset the costs of taking many small trades.

---

25 When $\zeta \neq 0$, $\gamma_1$ is a decreasing function of $\beta$, and the two are equal at approximately 0.5124.

26 This is true even if the optimal $\Theta$ does not become unbounded as $\beta \to \gamma_1$. Equation (21) is valid for any utility function, and when $\beta > \gamma_1$ many policy values $\Theta$ can give infinite utility so they cannot be ranked by expected utility.

27 Utility does not necessarily become unbounded in the absence of transaction costs if either the investor receives negative utility just from holding stocks at a paper loss or sales may be forced by exogenous events. In these cases, the optimal $\Theta$ may be strictly greater than 1 since selling right at the reference level creates the possibility of imminent disutility. These extensions are considered later in this paper.
Changes in $\mu$ have very little effect on size of optimal realized gains, $\Theta$, particularly for scaled-TK utility. The value function is strongly increasing in $\mu$, but this is due to the reduction in the average time before a sale occurs rather than any change in policy. The lower sales point, $\theta$, is affected more. For large or small $\mu$, the option to reset the reference point is less valuable than for intermediate values of $\mu$. It is less valuable for large $\mu$ because the asset price grows quickly enough for gains to be realized without a reset which necessitates paying the transactions costs. Conversely, with a low $\mu$, there is less value in resetting the reference level since future sales at a gain will take longer to be realized on average. So $\theta$ is highest for intermediate values of $\mu$. This is true in particular for modified-TK utility for which the option to reset is less valuable.

The standard deviation, $\sigma$, has even less effect on the optimal $\Theta$ than does $\mu$ — decreasing it by an amount imperceptible in the graphs. As with $\mu$, increasing $\sigma$ lowers the lower sales point, $\theta$. Again this can be explained with the option analogy. As shown in Figure 3, the value function decreases faster than the utility of a loss when $x$ falls, and the asset is sold when $v$ drops below the difference between the reset value and the disutility of the loss. The disutility of realizing a loss of a specific size does not depend on $\sigma$, but the option value of resetting the reference level is increasing in $\sigma$ because the continuation “payoff” is larger. Therefore, with higher $\sigma$, the asset value, and $v(x)$, must decrease by more before a reference-level-resetting sale is optimal.

The effects of reference level growth parameter, $\zeta$, can be seen through the reduced form equation. Increasing $\zeta$ is equivalent to decreasing $\mu$ by the same amount and decreasing $\delta$ by an amount $\beta$ times as large. These effects are mostly offsetting and no graph need be shown.

6. Model Extensions

In this section we examine several extensions to the realization utility model. We consider taxes, a stochastic reference level, and ongoing evaluation utility, and exogenously forced sales.

Taxes

Taxes are an obvious extension to the model since they affect the actual gains and losses from sales. In fact, capital gains are a near perfect fit with realization utility since they are typically due only upon sale of the asset. Here we analyze the effects of a simple tax code on realization utility. For simplicity, we assume that the investor’s reference level and his tax basis are equal. This means that the investor uses as his reference level the total purchase cost of the asset, as we have been assuming as one of our defaults; however, we do not necessarily assume that the investor takes taxes fully into account when assessing his return.

---

28 The maximum $\theta$ for a scaled-TK utility investor who ignores the transactions costs of the repurchase in assessing his gain ($\kappa = 1-k_s$) occurs at $\mu \approx 0.3\%$ and is not shown.

29 This assumption simplifies the analysis but precludes a growing or stochastic reference level and some of the interpretations of how transactions costs affect the reference level.
Suppose capital gains are taxed at a flat rate $\tau$ independent of the holding period. When an investor sells an asset that has a value of $X_n$ and a reference level and tax basis of $R_{n-1}$, he has a taxable gain of $X_n(1-k_x) - R_{n-1}$. After paying the tax and a second transaction cost to reinvest, he will have a net value and new cost basis and reference level of

$$R_n = \frac{X_n(1-k_x)(1-\tau) + \tau R_{n-1}}{1+k_p} = \frac{\Theta(1-k_x)(1-\tau) + \tau}{1+k_p} R_{n-1}. \quad (22)$$

The second equality holds for a constant-$\Theta$ sales policy. A similar relation with $\theta$ holds for a sale at a loss.

His utility burst depends on how he subjectively views the costs and taxes. If he ignores them completely, his gain is $\Delta x = (X_n - R_{n-1})/R_{n-1} = X_n/R_{n-1} - 1$, as before. At the opposite extreme, if he is fully cognizant of both, his gain is $R_n/R_{n-1} - 1$. If he ignores only the reinvestment cost, his gain is $1 - \frac{(1-k_x)}{1+k_p}(1-\tau)$. In general, the investor’s subjective gain can be expressed as $\kappa_1 X_n/R_{n-1} - \kappa_2$ with parameters adjusted to measure his subjective views.\(^{30}\)

Our original analysis can be applied directly with only minor changes. The utility burst is $e^{-\delta t} R^\beta u(\kappa_1 X/R - \kappa_2)$. The value function remains homogeneous of degree $\beta$ in $X$ and $R$, and still has a general solution of $C_1 x^{\gamma_1} + C_2 x^{\gamma_2}$ with the same exponents, $\gamma_1$ and $\gamma_2$. A constant proportional sales strategy will remain optimal. At a sale at either $\Theta R$ or $\theta R$, the boundary conditions are

$$\begin{align*}
v(\Theta) &= u(\kappa_1 \Theta - \kappa_2) + (1+k_p)^{-1}[\Theta(1-k_x)(1-\tau) + \tau]^{\beta} v(1) \\
v(\theta) &= u(\kappa_1 \theta - \kappa_2) + (1+k_p)^{-1}[\theta(1-k_x)(1-\tau) + \tau]^{\beta} v(1). \quad (23)
\end{align*}$$

These relations replace those in (12) and can be used to determine the unknown coefficients, $C_1$ and $C_2$ in terms of $\Theta$ and $\theta$ and then maximized or the smooth pasting conditions modified from (15) can be used.

The effect of taxes can be substantial. For example, adding a capital gains tax rate of 15% to the default parameters for scaled-TK\(^{31}\) changes the no sales region from (0.292, 1.052) to (0.066, 1.200) for an investor who recognizes the reinvestment cost in measuring his returns or from (0.417, 1.033) to (0.108, 1.188) for an investor who ignores it. There are similar effects for modified-TK utility.

It is perhaps surprising that receiving a tax refund on a loss does not make sales at losses more common. An expected utility maximizer understands that he has a loss even before it is

---

\(^{30}\) For the three cases mentioned. $\kappa_1 = \kappa_2 = 1$, $\kappa_1 = (1-k_x)(1-\tau)/(1+k_p)$ and $\kappa_2 = 1 - \tau/(1+k_p)$, or $\kappa_1 = (1-k_x)(1-\tau)$ and $\kappa_2 = 1-\tau$. Other interpretations are also possible as before.

\(^{31}\) The default parameters are $\mu = 10\%$, $\sigma = 30\%$, $k_x = k_p = 1\%$, $\alpha_G = \alpha_L = 0.5$, $\beta = 0.25$, $\lambda = 2$, $\delta = 4\%$, $\zeta = 5\%$. Since $\zeta$ must be zero to keep the reference level and the tax basis equal, we also reduced $\mu$ and $\delta$ to their effective levels of $\mu - \zeta = 5\%$ and $\delta - \zeta = 2.75\%$ for this comparison.
realized; he sells at a loss to receive the cash benefit of a tax refund. However, a realization utility investor avoids realizing losses doing so only to make gains more likely or larger in the future. Taxes do make him less reluctant to realize his loss due to the tax refund, but they also make after-tax gains smaller in the future. Since the optimally realized gain is smaller than the optimally realized loss, its marginal utility is higher, and taxes have a much larger impact on gains. This reduces the benefit of realizing losses gains to make gains more likely and so both types of sales are postponed.

Capital gains taxes can also induce strictly risk-averse realization-utility investors to sell at a loss. This is something he would never do in the absence of a capital gains tax because for concave utility the marginal utility of any loss exceeds that of any gain always so the gains cannot be split into smaller parts to give higher total utility than that from a loss of the same total size.

**Stochastic Reference Level**

In some settings the reference level might evolve stochastically over time. For instance, the investor might periodically learn information about what comparable investments are worth or might compare the return realized on the asset to the performance of the market or another index. If the reference level increases at the interest rate or some other rate, that rate might change randomly over time. Here we model the reference level as evolving randomly according to a lognormal process

\[
\frac{dS}{S} = \mu dt + \sigma_S d\omega_S, \quad \frac{dR}{R} = \zeta dt + \sigma_R d\omega_R, \quad \text{cov}[dS, dR] = \sigma_{SR} S R dt.
\]

Now Ito’s lemma and the law of iterated expectations give the value function as the solution to

\[
0 = \frac{1}{2} \sigma_S^2 X^2 Y_{xx} + \sigma_{SR}^2 X Y_{xR} + \frac{1}{2} \sigma_R^2 R^2 Y_{RR} + \mu X Y_X + \zeta R Y_R + Y_t.
\]

(25)

If the utility burst is \(R^\beta u(\kappa x - 1)\) as before, the value function remains homogeneous of degree \(\beta\), \(\bar{Y} = e^{-\delta t} R^\beta v(x)\), and (25) can be simplified to

\[
0 = \frac{1}{2} \sigma_x^2 x^2 v'' + (\mu_x - \zeta) xv' - (\delta - \zeta \beta)v
\]

where

\[
\delta_x \equiv \delta + \frac{1}{2} \beta (1 - \beta) \sigma_R^2, \quad \sigma_x^2 \equiv \sigma_S^2 - 2\sigma_{SR} + \sigma_R^2 = \sigma_S^2 + \sigma_R^2 (1 - 2\sigma_{SR}/\sigma_R^2)
\]

\[
\mu_x \equiv \mu + (1 - \beta) (\sigma_R^2 - \sigma_{SR}) = \mu + (1 - \beta) \sigma_R^2 (1 - \sigma_{SR}/\sigma_R^2).
\]

(26)

which is clearly identical to the base model with simple redefinitions of a few parameters. Since the sales points, \(\theta\) and \(\Theta\), are increasing in \(\mu\) and decreasing in \(\sigma\) and \(\delta\), a stochastic reference level will have mixed results depending on the parameter values.

Since \(\beta\) must lie between 0 and \(\alpha_G\), and \(\alpha_G < 1\), the effects of a stochastic reference level can be bounded. The maximum adjustment in \(\delta_x\) occurs for \(\beta = 0.5\), while for \(\beta = 0\) or 1, there is no adjustment. \(\delta_x\) always exceeds \(\delta\) but can do so by no more than 0.125\(\sigma_R^2\). As few indices have volatilities more than 30%, the largest adjustment to \(\delta_x\) is about one percentage point, and this
has little effect on optimal strategies. There is no adjustment to the mean, \( \mu \), if the stock’s beta with respect to the index, \( \sigma_{\text{SR}}/\sigma_{R}^2 \), is 1 which should be true on average for typical indices. Even for high and low beta stocks, the adjustment should be a few percentage points at most; however, this can have some effect on optimal strategies, particularly \( \theta \). The largest impact of a stochastic reference level probably comes through \( \sigma_x \). If the stock and the reference asset have equal variances and are uncorrelated, the adjusted standard deviation is 41% greater than \( \sigma_S \). This provides an upper bound on the magnitude of the effect.

If a realization-utility investor’s reference level is linked to the performance of an index in this fashion, then we should see an increase in volume after the asset has markedly outperformed or (assuming \( \theta > 0 \)) under-performed the index.

Another manner in which an investor might update his reference level is to modify it based on the recent price history of the asset. For example, even if the stock was selling for more than its initial purchase price, the investor may feel he is experiencing a loss if the price has fallen shortly before a sale. This feeling could be due to regret of having missed a favorable sales opportunity. Alternatively, if the stock price has been below the reference level for a while and then moves up, the investor might feel as if he has realized a gain even if the sale occurs at a price below the original purchase price.

To model this sentiment in a simple fashion, we assume the reference level is updated as an exponentially smoothed average of the asset position values since the previous purchase. The reference level is set to \( R_0 \) by a sale at time 0 and modified continually so that at time \( t \) it is \( R_t = e^{-\eta} R_0 + \eta \int_0^t e^{-\eta} X_{t-} ds \). The parameter \( \eta \) determines the relative importance of recent prices; the larger it is, the more important is the recent past. The average lag represented in the average at time \( t \) is \( (1 - e^{-\eta})/\eta \).

For this model, the evolution of the reference level is \( dR = \eta (X - R) dt \), and the value function equation is

\[
0 = \frac{1}{2} \sigma^2 X^2 Y_{XX} + \mu X Y_X + \eta (X - R) Y_R + Y_r. \tag{27}
\]

As shown in the Appendix, the solutions to this equation is

\[
C_1 x^{q/2} M(2q - 1, 1 \pm h, z) + C_2 x^{q/2} z^h M(2q - 1 \mp h, 1 \mp h, z)
\]

where

\[
z \equiv \frac{2 \eta x}{\sigma^2}, \quad q \equiv 1 - \frac{2(\mu + \eta)}{\sigma^2} \pm h, \quad h \equiv \sqrt{(2\mu + 2\eta - \sigma^2)^2 + 8\sigma^2(\delta + \beta \eta)}
\]

\[
M(a, b, z) \equiv \frac{\Gamma(b)}{\Gamma(a)} \sum_{n=0}^{\infty} \frac{\Gamma(a + n) z^n}{\Gamma(b + n) n!}.
\]

\( M(\cdot) \) is the confluent hypergeometric function of Kummer, and \( \Gamma(\cdot) \) is the gamma or generalized factorial function. Figure 9 plots the optimal sales points and the value function \( v(1) \) against the smoothing parameter \( \eta \) using the default parameters for scaled-TK utility.
There are two effects of this smoothing one dominating in the long term and the other in
the short term. With smoothing, even if the stock price suffers substantial losses, \( X \) cannot
remain too far below \( R \) in the long term. This means that profitable sales will occur often
and also that there is reduced incentive to sell at a loss to reset the reference level as the mean-
reversion takes care of this automatically. This effect is much stronger the larger is \( \eta \), and the
value function rises in value for high \( \eta \) while \( \theta \) drops quickly to 0. With a small \( \eta \) there is much
less mean reversion so profitable sales don’t occur as frequently as \( X \) can fall much further below
\( R \). That means the incentive to sell at a loss and reset the reference level remains.

This extension models an investor feeling as if he has realized a gain by selling a stock
when its price rises after it has been depressed for some time. This is consistent with a suggestion
made by Odean (1999).32

Ongoing Utility

We have assumed so far that investors obtain utility only directly from the sale of an asset
at a gain or a loss. However, once we’ve put aside the EUT view that utility only arises from
spending on consumption and allow it to be realized simply from selling an asset without spend-
ing the proceeds other than to reinvest, the notion that holding an asset with a paper gain might
also provide utility seems an obvious extension.

We assume now that an investor realizes utility as described previously when selling an
asset, but also obtains a continuous utility flow, \( \Psi(X, R, t) \), from holding an asset with a paper
gain or loss. Now equation (7) for lifetime expected utility is modified to

\[
0 = \frac{\mathbb{E}[dY(X, R, t) + \Psi dt]}{dt} = \mu X Y_X + \frac{1}{2} \sigma^2 X^2 Y_{XX} + \zeta Y_R + Y + \Psi(X, R, t) .
\]  

(29)

This equation is reminiscent of the Bellman equation in the standard consumption-portfolio
problem with \( \Psi \) serving the role of the utility of consumption. The main difference is that there
is no decrease in wealth needed to provide this utility.

If this utility has the same reference-level scaling as the direct utility function, \( \Psi(X, R, t) = e^{-\delta t} R^\rho \psi(x) \), we can use still use the homogeneity property to write the reduced valuation
equation. As a simple example, suppose ongoing utility has the same functional form as utility
from a sale, \( \psi(x) = \hat{\rho}u(\hat{\kappa} x - 1) \). The parameter \( \hat{\rho} \) measures the relative importance of paper
profits compared to realized profits. For example, if \( \hat{\rho} = 0.1 \), then the investor receives as much
utility from selling a stock with a $25 dollar gain as he does from holding that stock for ten years
with an average paper profit of $25 (apart from discounting). \( \hat{\kappa} \) is the subjective transactions cost
parameter which may differ from that used for realized profits. Unless the investor thinks in
terms of “after-cost” paper profits, \( \hat{\kappa} \) would be 1. If he does think of paper profits as potential

32 Figure 6 in Odean (1999) confirms that investors tend to sell stocks with both paper gains and losses if their prices
have risen rapidly for the past 20 days.

33 Assuming that ongoing and sale utility have the same functional form is merely a convenience and a reasonable
presumption. It is not required to solve the model.
after-cost actual profits, \( \hat{k} \) could be \( 1 - k_s \) or \( K \) or another number just as in the original analysis. The reduced form equation is

\[
0 = \frac{1}{2} \sigma^2 x^2 v'' + (\mu - \zeta) xv' - (\delta - \zeta \beta) v + \hat{\rho} u (\hat{k} x - 1) . \tag{30}
\]

The differential equation is the same as before apart from the inhomogeneous term requiring a particular solution. This term differs in the gain and loss regions, so we need to solve each region separately. The gain-loss region is divided at \( x = 1/\hat{k} \). The general solution is derived in the Appendix

\[
v(x) = \begin{cases} C_{G_1} x^{\gamma_1} + C_{G_2} x^{\gamma_2} + v_G(x) & x > \hat{k}^{-1} \\ C_{L_1} x^{\gamma_1} + C_{L_2} x^{\gamma_2} + v_L(x) & x \leq \hat{k}^{-1} \end{cases}
\tag{31}
\]

where \( \gamma_1 \) and \( \gamma_2 \) are defined as before.

For modified-TK the particular solutions are

\[
v_G(x) = -\frac{\hat{\rho}}{(\delta - \zeta \beta) \alpha_G} - \frac{\hat{\rho} \hat{k}^{\alpha_G}}{\alpha_G [\frac{1}{2} \sigma^2 (\alpha_G - 1) + (\mu - \zeta) \alpha_G - (\delta - \zeta \beta)]} x^{\alpha_G} \\
v_L(x) = -\frac{\lambda \hat{\rho}}{(\delta - \zeta \beta) \alpha_L} - \frac{\lambda \hat{\rho} \hat{k}^{\alpha_L}}{\alpha_L [\frac{1}{2} \sigma^2 (\alpha_L - 1) + (\mu - \zeta) \alpha_L - (\delta - \zeta \beta)]} x^{\alpha_L} . \tag{32}
\]

For scaled-TK utility the particular solutions are

\[
v_G(x) = 2 \sigma^{-2} \hat{\rho} (\hat{k} x)^{\gamma_2} B\left((\hat{k} x)^{-1}; \gamma_2, \gamma_2 - \alpha_G, \alpha_G + 1\right) - (\hat{k} x)^{\gamma_2} B\left((\hat{k} x)^{-1}; \gamma_1, \gamma_2 - \alpha_G, \alpha_G + 1\right) \\
v_L(x) = 2 \sigma^{-2} \lambda \hat{\rho} (\hat{k} x)^{\gamma_2} B\left((\hat{k} x)^{-1}; \gamma_1, \gamma_2, \alpha_L + 1\right) - (\hat{k} x)^{\gamma_2} B\left((\hat{k} x)^{-1}; \gamma_1, \gamma_2, \alpha_L + 1\right) \tag{33}
\]

where \( B(z; a, b) \) is the incomplete beta function.\(^{34}\)

The four \( C \) coefficients remain to be determined as well as the optimal strategies \( \theta \) and \( \Theta \). As before, the boundary conditions at \( \theta \) and \( \Theta \) supply two equations. Two more equations come from the continuity of the value function and its derivative at the gain-loss boundary \( x = 1/\hat{k} \).

The value function must be continuous by the law of iterated expectations since it is a martingale whose expected change must be zero; furthermore, since \( x \) is a diffusion, the derivative must also be continuous for the same reason. The four equations are\(^{35}\)

\(^{34}\) The incomplete beta function is defined as \( B(z; a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt \). See Abramowitz and Stegun (1970). For the second incomplete beta function in \( v_G \), the integral diverges as \( x \to \infty \) since \( \gamma_2 - \alpha_G < 0 \); however, in this case, \( (\hat{k} x)^{\gamma_2} \to 0 \), and the product remains finite. Similarly, for the first incomplete beta function in \( v_L \), the integral diverges as \( x \to 0 \) since \( \gamma_1 > 0 \); however, in this case, \( (\hat{k} x)^{\gamma_1} \to 0 \), and the product remains finite.

\(^{35}\) Immediately after a sale and repurchase, \( x = 1 \) which is in the loss region; so \( v(1) = C_{L_1} + C_{L_2} + v_{L}(1) \).
The optimal sales policy can be determined either by numerically maximizing \( v(1) = C_{L1} + C_{L2} + v_L(1) \) or applying the smooth pasting conditions at the sales points

\[
\begin{align*}
C_{L1} \Theta^\gamma + C_{L2} \Theta^{\gamma_2} + v_L(\Theta) &= u(\kappa \Theta - 1) + (K \Theta) \beta [C_{L1} + C_{L2} + v_L(1)] \\
(C_{L1} - C_{L2}) \hat{k}^{-\gamma_1} + (C_{L2} - C_L) \hat{k}^{-\gamma_2} + v_L(\hat{k}^{-1}) &= 0 \\
(C_{L1} - C_{L2}) \gamma_1 \hat{k}^{-\gamma_1} + (C_{L2} - C_L) \gamma_2 \hat{k}^{-\gamma_2} + v_L(\hat{k}^{-1}) &= 0.
\end{align*}
\] (34)

Figure 10 shows the optimal policies and value function for this model for the default parameters for modified-TK utility. The most salient feature is that there are two local maxima, one with a decreasing value function (illustrated in red) and the other with an increasing value function (in blue). The global maximum switches from the decreasing-\( \nu \) maximum to the increasing-\( \nu \) maximum when \( \hat{\rho} \) reaches a bit over 0.13. For both maxima, \( \Theta \) increases with \( \hat{\rho} \) because the investor’s reluctance to realize a loss is reduced by the increasing incentive to eliminate the ongoing negative utility which is proportional to \( \hat{\rho} \). For the decreasing-\( \nu \) maximum, \( \Theta \) is about 1.1 and is decreasing slightly with \( \hat{\rho} \) as shown; however, for \( \beta = 0 \) (not shown), \( \Theta \) increases with \( \hat{\rho} \). As \( \hat{\rho} \) increases the investor is more reluctant to sell at a gain and give up the ongoing positive utility. However, with \( \beta = 0.25 \) this is offset by the increased value when resetting the reference point. For the increasing-\( \nu \) maximum, \( \Theta \) is extremely high. For \( \beta = 0 \), \( \Theta \) is infinite and the investor never realizes a gain; \( \nu \) is just the present value of the ongoing utility. For \( \beta = 0.25 \), the investor eventually sells to realize a gain after the stock has appreciated, and the reset value \( (K \Theta)^\beta \) becomes enormous.\(^{36}\)

There is one other distinction of this model. In the original model, the investor could avoid negative utility bursts entirely if desired. He only realized losses to reset the reference level and make future gains more likely. But in this model, negative utility outcomes cannot be absolutely precluded as the investor might experience paper losses. Therefore, if the environment is unfavorable, the investor may not enter the market at all. He will not participate if the initial value does not exceed his reservation level. An obvious choice for the reservation level is 0 which is the utility of no gains and losses, though it could be higher or lower if the investor had other alternatives or had to choose from a set of distasteful possibilities.\(^{37}\)

\[^{36}\] This \( \Theta \) is never less than 395 in the graph and is economically meaningless. While such extreme asset growth might be observed on occasion, it is beyond the realm of possibility that this or any economic model can make such precise extrapolations.

\[^{37}\] For example, suppose the investor must choose an asset. If a risk-free asset is available, the utility of that option makes a sensible candidate as the reservation level utility. If the interest rate is \( r \) and the reference level is assumed to increase at \( \zeta = r \), then the risk free asset generates zero ongoing utility and, to avoid transactions costs, will never be sold. In this case the reservation utility is still 0.
Figure 11 shows typical participation constraints for different values of $\hat{\rho}$ and a minimum participation value of $v(1) = 0$. An investor with a given value of $\hat{\rho}$ is willing to invest in an asset whose mean and variance lie at any point above the relevant curve. For $\hat{\rho} = 0$, there is no participation constraint since the investor can always refrain from realizing any losses. Consistent with this, generally, the greater is the relative importance of ongoing utility, the higher is the participation constraint; however, we can see from the graph that this is not universally true. The solid portions of the $\hat{\rho} = 0.04$ and $\hat{\rho} = 0.06$ participation constraints indicate when the optimal strategy is to never realize a gain ($\Theta = \infty$). For the $\hat{\rho} = 0.02$ constraint, the investor never realizes a loss ($\theta = 0$). The constraints are not very tight, though an investor will not necessarily invest in an asset simply because it meets the participation constraint. He will choose that asset or those assets which provide the highest utility.

The participation constraints have the same inverted U shape as indifference curves for low $v$ because that is exactly what they are. For larger values of $\hat{\rho}$ the participation constraints are higher and flatter. That is, for a given $\sigma$, an investor with a higher $\hat{\rho}$ demands a higher $\mu$ to willingly purchase the asset. The higher the stock’s volatility, the more likely is it to decrease in price resulting in paper losses and negative ongoing utility.

**Exogenous Events**

As a final extension to our model we consider exogenous realization events — investors may have a change in utility for reasons other than the voluntary and utility-maximizing sale of their positions. The standard way to model exogenous events is using a Poisson shock with a specified intensity, $\rho$. With Poisson events, (7) is modified to

$$0 = \mathbb{E}[dY(X, R, t) + \rho dt[\Psi(X, R, t) - Y(X, R, t)]]$$

$$= \mu X Y'_x + \frac{1}{2} \sigma^2 X^2 Y''_x + \zeta Y'_x + Y'_t + \rho[\Psi(X, R, t) - Y].$$

The first four terms are the continuous change in the value function driven by the diffusion of $X$ and any growth in $R$. The final term is the change caused by the Poisson event. The value function changes discontinuously from $Y$ to $\Psi$ with probability $\rho$ per unit time.

There are a number of economic interpretations of the function, $\Psi(X, R, t)$. The simplest is $\Psi \equiv 0$ corresponding to the assumption that the investor is forced out of the market by the Poisson event with no final utility burst. This could model the investor’s death or simply his disinterest in actively following his investment any further. When $\Psi = 0$, the reduced form of the valuation equation equivalent to (10) is

$$0 = \frac{1}{2} \sigma^2 x^2 v'' + (\mu - \zeta) x v' - (\rho + \delta - \zeta \beta) v.$$

The Poisson-driven shortening of the horizon is mathematically equivalent to adding the intensity parameter $\rho$ to the discount factor $\delta$ so our original analysis can be applied directly by just modifying that parameter. Typically this change has little effect as shown in Figures 7 and 8. The exception is when a transversality condition comes into play. Increasing $\delta$ loosens each of
the three transversality conditions which can have a dramatic impact.

There is a much larger effect if the Poisson shock that forces the investor out of the market also generates one final utility burst from a sale. This is the model studied by Barberis and Xiong (2012). For this model, \( \Psi = e^{-\delta t} R^0 u(\hat{\kappa} x - 1) \) is the utility received from the final sale. We denote the subjective adjustment here by \( \hat{\kappa} \) to allow for the possibility that the adjustment for a final forced sale differs from that for a voluntary sale.\(^\text{38}\) Using the homogeneity property the reduced valuation equation is

\[
0 = \frac{1}{2} \sigma^2 x^2 v^* + (\mu - \zeta) x v' - (\mu + \delta - \zeta \beta) v + \rho u(\hat{\kappa} x - 1). \tag{38}
\]

This equation is identical to (30) with \( \delta \) replaced by \( \delta + \rho \) and \( \hat{\rho} = \rho \). All the analysis and interpretations given there hold here subject to the reinterpretation of the parameters. In particular the subjective discount rate is higher.

Figure 12 shows the participation constraint for an investor facing forced Poisson exits. They are tighter than those for ongoing utility with the same parameter. The intuition is simple, the investor is less willing to invest if he might be forced to stop at a loss with no chance of future gains.

7. Model Predictions

Our paper makes several direct predictions about trading activity and indirect predictions about asset pricing. Our various models indicate specific, though parameter dependent, forecasts about the timing of trades and the prices at which they will be executed. On the other hand, our predictions about equilibrium prices are only indirect since there is no equilibrium model in our paper.

Asset Pricing

If the typical realization utility investor has a low \( \beta \), a high \( \zeta \), and a high \( \delta \), then as discussed above, he prefers stocks with a higher volatility for any given expected return in contrast to the standard mean-variance model. Even for moderate values of \( \beta \) and \( \zeta \), he has more tolerance for volatility than does the typical mean-variance investor because his indifference curves have a shallower slope than the capital market line.

Now consider an economy dominated by mean-variance investors with some realization utility investors. The former will hold mean-variance efficient portfolios, but the presence of the latter will mean there is excess demand for high-mean, high-variance stocks. A stock with a high beta will provide both, but to the extent that the mean return is explained by beta, but there will

\(^{38}\) In particular, if the Poisson event forces out of the market an investor who typically fully accounts for transactions costs, i.e., \( \kappa = K = (1-k_s)/(1+k_p) \), then the adjustment for the final sale would not include the repurchase transactions cost; i.e., \( \hat{\kappa} = (1-k_s) \).
also be excess demand for high residual risk stocks which mean-variance investors will only hold in diversified portfolios.

In keeping with the standard CAPM framework, suppose we can analyze the equilibrium in a one-period framework in which all investors have homogeneous beliefs, the mean-variance investors hold the tangency portfolio, \( t \), and the realization utility investors hold some other portfolio \( p \) which represents the fraction \( w \) of the entire market.\(^\text{39}\) We know that the tangency portfolio predicts expected returns in the usual way, and since the market portfolio is a combination of it and portfolio \( p \), we can determine the relation between the market portfolio and expected returns.

In particular, for the tangency portfolio

\[
\frac{\mu_t - r}{\mu_m - r} = \frac{\sigma_t^2}{\sigma_m^2} = \frac{\beta_{mt}}{\rho_{mt}^2} > \beta_{mt}
\]

(39)

where \( \beta_{mt} \) is the market beta of the tangency portfolio and \( \rho_{mt} \) is the correlation between the returns on the market and tangency portfolios. This correlation must be less than one (unless portfolio \( p \) is exactly the market mixture) so the expected rate of on the tangency portfolio exceeds the prediction of the market-based CAPM. This means that on average the stocks underrepresented in portfolio \( p \), and therefore overrepresented in the tangency portfolio, will appear to have positive market alphas.

For any stock or portfolio \( j \)

\[
\mu_j - r = \frac{\sigma_j^2}{\sigma_t^2}(\mu_t - r) \quad \mu_m - r = w(\mu_p - r) + (1 - w)(\mu_t - r)
\]

\[
\sigma_{jm} = w\sigma_{jp} + (1 - w)\sigma_{jt} \quad \sigma_m^2 = w^2\sigma_p^2 + 2w(1 - w)\sigma_{pt} + (1 - w)^2\sigma_t^2
\]

(40)

imply for portfolio \( p \)

\[
\frac{\mu_p - r}{\mu_m - r} = \beta_{mp} - \frac{w\sigma_p^2}{\sigma_m^2(1 - w\beta_{mp})}
\]

(41)

where \( \beta_{mp} \) is the market beta of the portfolio of stocks held by the realization utility investors and \( s_p^2 \) is its residual variance.

This derivation is completely general. The market can always be decomposed into the tangency portfolio and some residual portfolio \( p \). The important assumptions are that mean-variance investors predominate in the market which assures us that \( w \) is small and that the other investors are realization utility investors which helps us identify what portfolio \( p \) might contain.

\(^{39}\) Investors may exhibit both traits. For example, an investor might hold his retirement account in a mean-variance efficient portfolio, but trade on his personal account according to realization utility. Portfolio \( p \) represents that portion of their wealth held outside of the tangency portfolio, and \( w \) is the fraction of market wealth held in portfolio \( p \).
We can expand the denominator in a three term Taylor expansion using the exact form of the mean-value theorem as

\[
\left( \frac{\mu_p - r}{\mu_m - r} \right) = \beta_{mp} - \frac{w s_p^2}{\sigma_m^2} \left( 1 + w \beta_{mp} + \overline{w}^2 \beta_{mp}^2 \right) = \beta_{mp} \left( 1 - \frac{w s_p^2}{\sigma_m^2} \right) - \frac{w s_p^2}{\sigma_m^2} - \beta_{mp} \frac{w \overline{w}^2 s_p^2}{\sigma_m^2}.
\]  

(42)

where \( \overline{w} \in (0, w) \). The relation between portfolio \( p \) and its beta is shallower than predicted by the CAPM. In addition the relation should appear concave and apparent risk premium for its residual risk will be negative.

We should expect these properties to carry through to the individual assets in \( p \). For any individual stock \( j \), the general pricing relation is

\[
\left( \frac{\mu_j - r}{\mu_m - r} \right) = \frac{\beta_{mj} - \frac{w s_j}{\sigma_m^2}}{1 - w \beta_{mp}}.
\]  

(43)

The covariance \( \sigma_{jp} \) can be expressed as \( \sigma_{jp} = \beta_{mj} \beta_{mp} \sigma_m^2 + s_{jp} \) where \( s_{jp} \) is the covariance between the idiosyncratic risks of asset \( j \) and portfolio \( p \). Therefore

\[
\left( \frac{\mu_j - r}{\mu_m - r} \right) = \beta_{mj} - \frac{w s_{jp}}{(1 - w \beta_{mp}) \sigma_m^2} = \beta_{mj} - \frac{w s_{jp}}{(1 + w \beta_{mp} + \overline{w}^2 \beta_{mp}^2)}.
\]  

(44)

This relation shows that a stock whose residual risk covaries positively with the residual risk of portfolio \( p \), will have an expected rate of return smaller than predicted by the CAPM and further that the relation will depend on the stock’s residual risk and its beta in a non-linear fashion.

Stocks held by the realization utility investors, and therefore that are in portfolio \( p \), would tend to have the largest residual covariances and be most affected. The size of the effect will also depend on \( w \); that is, on how important realization utility investors are in the market.

Fama and MacBeth (1973) found

\[
\hat{R}_p = \gamma_0 + \gamma_1 \times \hat{\beta}_{mp} + \gamma_2 \times \hat{\beta}_{mp}^2 + \hat{\epsilon}_p
\]  

(45)

\[
\hat{R}_p = \gamma_0 + \gamma_1 \times \hat{\beta}_{mp} + \gamma_2 \times \hat{\beta}_{mp}^2 + \gamma_3 \times \hat{s} + \hat{\epsilon}_p
\]

which agrees with the prediction that the square of the market beta should depress average returns. Their result on residual risk is not in agreement with our model’s predictions. However, more recently, Ang, Hodrick, Xing, and Zhang (2006) tested the CAPM and Fama-French 3-factor alphas for stocks with different total and residual variation. They found that the difference between the alphas of the highest (8.30%) and lowest (3.71%) volatility stocks was
−1.35% (t = −4.62) for the CAPM and −1.19 (t = −5.92) for FF-3. For residual risk they found the difference in alphas between the highest (8.16%) and lowest (3.83%) residual volatility stocks was −1.38% (t = −4.56) for the CAPM and −1.31 (t = −7.00) for FF-3.

Trading Probabilities and Volume

One of the predictions of our model is that realization utility investors make frequent trades, particularly to realize gains. To determine the trading frequency and volume, we compute the probabilities of trading at a gain or a loss as a function of holding time. We concentrate here on the voluntary sales since the Poisson-forced sales simply add constant trading volume and would affect any type of investor. In keeping with our previous analysis we assume that the asset price evolves according to a lognormal diffusion.

We treat one investment episode from initial purchase through ultimate sale as a diffusion process with two absorbing barriers, one when the stock price rises to $\Theta R$ and the other when it falls to $0 R$. To simplify the exposition, we work in log space with $z \equiv \ln(X/R)$. The evolution of $z$ is $dz = \mu dt + \sigma \omega$ with $\mu = \mu - \zeta - \frac{1}{2} \sigma^2$. The stochastic process starts at $z_0 = 0$; the absorbing barriers are $a \equiv \ln \theta$ and $b \equiv \ln \Theta$. The cumulative probabilities of the first hitting times at either barrier (without first hitting the other) are for $\mu' \geq 0$:

$$Q_0(t) = Q_0^\infty + \sum_{n=-\infty}^{\infty} \exp\left(\frac{\mu' z_n^*}{\sigma^2}\right) \Phi\left(\frac{b - z_n^* - \mu' t}{\sigma \sqrt{t}}\right) - \exp\left(\frac{\mu' z_n^*}{\sigma^2}\right) \Phi\left(\frac{b - z_n^* - \mu' t}{\sigma \sqrt{t}}\right)$$

$$Q_0(t) = Q_0^\infty + \sum_{n=-\infty}^{\infty} \exp\left(\frac{\mu' z_n'}{\sigma^2}\right) \Phi\left(\frac{a - z_n' - \mu' t}{\sigma \sqrt{t}}\right) - \exp\left(\frac{\mu' z_n'^*}{\sigma^2}\right) \Phi\left(\frac{a - z_n'^* - \mu' t}{\sigma \sqrt{t}}\right)$$

where $z_n' \equiv 2n(b - a)$ and $z_n^* \equiv 2b - z_n'$.

and $Q_0^\infty = \frac{1 - \exp(2\mu'a/\sigma^2)}{1 - \exp(-2\mu'(b-a)/\sigma^2)}$ and $Q_0^\infty = 1 - Q_0^\infty = \frac{\exp(2\mu'b/\sigma^2) - 1}{\exp[2\mu'(b-a)/\sigma^2] - 1}$

are the probabilities that the stock is eventually sold at a gain or loss, respectively.

These two formulae are valid for all of the models presented in this paper except the

---

40 If $\rho \neq 0$, then trading also occurs when the Poisson event arrives. To characterize sales at gains and losses we need to properly count the paper gains and losses that are associated with Poisson arrival. More detailed discussion is in our companion paper, Ingersoll and Jin (2012).

41 For $\mu' < 0$, the cumulative probabilities are

$$Q_0(t) = Q_0^\infty + \sum_{n=-\infty}^{\infty} \exp\left(-\frac{\mu' z_n^*}{\sigma^2}\right) \Phi\left(-\frac{\mu' t - b - z_n^*}{\sigma \sqrt{t}}\right) - \exp\left(-\frac{\mu' z_n^*}{\sigma^2}\right) \Phi\left(-\frac{\mu' t - b - z_n^*}{\sigma \sqrt{t}}\right)$$

$$Q_0(t) = Q_0^\infty + \sum_{n=-\infty}^{\infty} \exp\left(-\frac{\mu' z_n'^*}{\sigma^2}\right) \Phi\left(-\frac{\mu' t - a - z_n'^*}{\sigma \sqrt{t}}\right) - \exp\left(-\frac{\mu' z_n'^*}{\sigma^2}\right) \Phi\left(-\frac{\mu' t - a - z_n'^*}{\sigma \sqrt{t}}\right).$$

For $\mu' = 0$, either formula can be used; the constants are $Q_0^\infty = -a/(b-a)$ and $Q_0^\infty = b/(b-a)$. 

28
stochastic reference level extensions. When the reference level is an index or other correlated
lognormal price, the formula is valid with $\mu' = \mu - \zeta - \frac{1}{2} \sigma_z^2$ and $\sigma^2 = \sigma_z^2$ as defined in (26). For
the smoothed reference level described in (28), the probabilities have a different form. The
probabilities that the stock is eventually sold at a gain and loss are

$$Q_\phi^\infty \equiv \frac{P(A+1,B) - P(A+1,B\theta)}{P(A+1,B) - P(A+1,B\theta)} \quad \text{and} \quad Q_0^\infty = 1 - Q_\phi^\infty \equiv \frac{P(A+1,B\theta) - P(A+1,B)}{P(A+1,B) - P(A+1,B\theta)}$$

where $A \equiv 2 \frac{\mu + \eta}{\sigma^2}$, $B \equiv \frac{2\eta}{\sigma^2}$, and

$$P(a,z) = \frac{1}{\Gamma(a)} \int_0^z e^{-t} t^{a-1} dt$$

is the incomplete gamma function. Of course for each model, the relation between the sales
points and the various parameters differs.

Table I gives some representative values for the sales policies. The frequency of sales at
gains exceeds that of sales at losses by a factor of at least 10 for all intervals. For short periods,
the ratio is even larger.

<table>
<thead>
<tr>
<th>time</th>
<th>scaled-TK Utility</th>
<th>modified-TK utility</th>
<th>modified-TK utility with ongoing utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales points</td>
<td>$Q_\phi$</td>
<td>$Q_0$</td>
<td>$Q_\phi$</td>
</tr>
<tr>
<td>(\Theta 0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 month</td>
<td>1.033</td>
<td>0.417</td>
<td>1.110</td>
</tr>
<tr>
<td>2 months</td>
<td>71.08%</td>
<td>0.00%</td>
<td>22.90%</td>
</tr>
<tr>
<td>3 months</td>
<td>79.37%</td>
<td>0.00%</td>
<td>39.60%</td>
</tr>
<tr>
<td>6 months</td>
<td>83.12%</td>
<td>0.00%</td>
<td>48.90%</td>
</tr>
<tr>
<td>1 year</td>
<td>91.60%</td>
<td>0.17%</td>
<td>73.19%</td>
</tr>
<tr>
<td>2 years</td>
<td>94.11%</td>
<td>1.18%</td>
<td>81.02%</td>
</tr>
<tr>
<td>5 years</td>
<td>96.14%</td>
<td>2.93%</td>
<td>88.13%</td>
</tr>
<tr>
<td>10 years</td>
<td>96.58%</td>
<td>3.36%</td>
<td>91.64%</td>
</tr>
<tr>
<td>15 years</td>
<td>96.61%</td>
<td>3.39%</td>
<td>92.86%</td>
</tr>
<tr>
<td>20 years</td>
<td>96.61%</td>
<td>3.39%</td>
<td>93.32%</td>
</tr>
<tr>
<td>\infty</td>
<td>96.61%</td>
<td>3.39%</td>
<td>93.60%</td>
</tr>
</tbody>
</table>

$Q_\phi(t)$ and $Q_0(t)$ denote the probabilities that the stock is sold at a gain or loss, respectively,
within $t$ years of its purchase. The default parameters are used: $\mu = 10\%$, $\sigma = 30\%$, $k_s = k_p = 1\%$, $\beta = 0.25$, $\lambda = 2$, $\delta = 4\%$. For scaled-TK utility $\alpha_G = \alpha_L = 0.5$. For modified-TK utility $\alpha_G = -3$, $\alpha_L = 5$. The investor ignores the reinvestment cost in assessing his realized
gains ($\kappa = 1-k_s$).
Figure 13 plots the two trading probabilities, $Q_\theta(1)$ and $Q_\Theta(1)$, against $\mu$ and $\sigma$ for $k_s = k_p = 1\%$, $\alpha_G = -3$, $\alpha_L = 8$, $\beta = 0.25$, $\lambda = 1.5$, $\zeta = 5\%$, $\delta = 4\%$, $\rho = 0$ for an investor with modified-TK utility. The solid lines show the probabilities that the stock reaches the trading points, $\theta$ or $\Theta$, optimal for the given value of $\mu$ or $\sigma$ within one year of the initial purchase. The dotted lines show the probabilities of reaching fixed trading points — those that are optimal for $\mu = 10\%$ and $\sigma = 50\%$.

Increasing $\mu$ obviously makes hitting a fixed $\theta$ and $\Theta$ less and more likely, respectively, as shown by the dotted lines in Figure 13. This is obviously a mechanical effect of the stochastic process. But as $\mu$ increases $\Theta$ also gets larger. This makes the actual $Q_\theta$'s relation to $\mu$ as shown in the solid line start at a higher value for low $\mu$ and remain flatter than the fixed-$\Theta$ relation as $\mu$ increases. Similarly since $\theta$ is also increasing for low $\mu$, the actual $Q_0$ starts lower and remains flatter than the fixed-$\theta$ probability for most of the range of $\mu$. At high $\mu$, $\theta$ begins to fall again making the $Q_0$ curve decline more quickly that the fixed-$\theta$ curve.

As $\sigma$ increases, the cumulative probabilities of hitting a fixed sales point increase. This again is purely an effect of the stochastic process; the more variation there is in the process, the more likely is it that a fluctuation will push the stock price to one of the sales points quickly. As seen in Figure 8, the optimal $\theta$ and $\Theta$ both fall with an increasing $\sigma$ so the probabilities of reaching the optimal $\theta$ rises less rapidly while the probability of reaching the optimal $\Theta$ rises faster. In addition, each of these has an amplifying effect on the other. Once either $\theta$ or $\Theta$ is reached the other cannot be during this investment episode. Since, with higher $\sigma$, it is less likely that $\theta$ will be reached, it becomes more likely that $\Theta$ will be reached and vice versa. The total probability of any sale, $Q_\theta + Q_\Theta$, behaves more like that of a sale at a loss because $\Theta$ is much closer to the starting price than $\theta$ so volatility has less of an effect on its probability.

This analysis is consistent with the disposition effect, namely that investors have a higher propensity to sell stocks at gains than at losses. However, as shown in Table I, the probability of a sale at a gain is substantially larger than the probability of a sale at a loss. Furthermore, since $\Theta$ is typically much closer to 1 than is $\theta$, on average fewer stocks with paper gains would be held. This makes our results “too good” for the disposition effect using the measure proposed in Odean (1998).

Our model also makes a number of other predictions related to trading volume. Volume is higher in bull markets than in bear markets since $\Theta$ is typically closer to the reference level than $\theta$. A more subtle prediction is that volume should be highest in extreme bull markets when stock prices rapidly rise through optimal sales levels. That is, with proper empirical measures,

---

$^{42}$ Since $\Theta > e^{\mu - \zeta}$, the probability of reaching $\Theta$ within a year is zero when $\sigma = 0$. If $\Theta$ were smaller, then $Q_\Theta(1)$ would be 100% and that curve would initially be decreasing in $\sigma$.

$^{43}$ More detailed discussion is in our companion paper, Ingersoll and Jin (2012). We have not included here any Poisson-event “forced” sales which would tend to reduce the ratio of sales at gains and sales at losses. We have also ignored the distinction between subjective gains and losses and measured gains and losses which arises when the reference level is changing.
our model can link trading volume to the speed at which markets rise or fall. Volume will be lowest, not in bear markets, but rather in flat markets which do not move much in either direction. Trading volume will also be path dependent; we should expect volume to increase when a stock price reaches new highs or highs it has not achieved in a while. To a lesser extent this will be true of new lows as well. These effects should be most pronounced in the stocks most favored by realization utility investors.

Since realization utility investors concentrate in high beta and high residual risk stocks, volume should be higher for these stocks. Of course, institutional investors tend to hold the larger cap stocks, and their trading would probably offset this effect. On the other hand, our analysis of alphas showed that these stocks also should have negative alphas, so volume should be higher in overpriced stocks. There is less reason to suppose that institutional investors’ trading will offset this relation.

8. Concluding Remarks

In this paper, we discuss some desiderata for cumulative prospect theory in an intertemporal setting. We compare the distinctions and contrasts with expected utility theory. We analyze the properties of two utility forms, scaled-TK and modified-TK utility, and then use them to build a realization utility model in continuous time to study investors’ trading behavior. We also explore some extensions to the model examining taxes, a stochastic reference level, ongoing utility, and exogenously forced sales. Finally, we discuss some empirical implications of the model.

One direction for future research is to relax the assumption that realization-utility investors hold a single stock and solve a full portfolio problem. We do know that some diversification is optimal, but solving the portfolio problem adds several complications. Since diversification is optimal, some rebalancing must also be desirable, but this means that the optimal sales strategy for a stock depends not only on its reference level, but also on the prices and reference levels of all the other assets held in the portfolio. In addition, assumptions must then be made about the proper reference level for a position that consists of shares purchased for different prices at different times.

More fundamentally, we need to understand how the utility of gains and losses on separate assets or at different times are aggregated. Do a simultaneous gain and loss offset each other with utility assessed on the net gain or does narrow framing apply so the utility of each is determined separately and summed? Both assumptions are psychologically plausible, but either creates problems in multi-period models, particularly continuous-time models. If the investor’s utility is defined on the net gain aggregated across several assets, then when stocks move down together, he can sell simultaneously, and when they move up together, he can postpone realizing one gain by \( dt \). This strategy increases his total utility because the marginal utility of each of the

\[ W_0 = \sum_i (w_i W_0)^{\beta} v_i(1). \]

where \( v_i \) is the separate valuation of asset \( i \). For \( 0 < \beta < 1 \), this achieves a maximum with \( w_i \propto \{v_i(1)\}^{1/(\beta-1)}. \) This strategy typically would not be optimal without the reinvestment restriction, but it demonstrate that investing in a single stock is dominated so some diversification must be optimal.

\[ W_0 = \sum_i (w_i W_0)^{\beta} v_i(1). \]
separate gains is larger while the marginal utility of the merged loss is smaller. On the other hand, if narrow framing applies across time even for individual assets, then the investor should conclude an investing episode by realizing the gain continuously over some interval of time again relying on the utility function’s concavity to increase his total utility. While these policies are formally ideal, they wouldn’t seem to be economically meaningful unless we believe that investors do play such mind games on themselves.

It is also worthwhile mentioning that it is a non-trivial task to incorporate cumulative probability weighting into an intertemporal setting because law of iterated expectation would not hold if probabilities are transferred into decision weights, and therefore martingale price will break. As shown in Barberis (2011) and Ingersoll (2011), cumulative probability weighting typically induces time inconsistency and certain rules that define economic actions when irrationality arises need to be imposed in order to further model this type of behavior.

Another related research topic is a deeper analysis on the connection between realization utility and the well-known disposition effect. How the trading patterns and volume evolve over time in different markets can also be further investigated. Other applications such as the relation between realization utility and momentum are worthwhile studying. A more detailed discussion can be found in Ingersoll and Jin (2012).
Appendices

A: Proof of Proposition 1

To prove this theorem, we use the following lemma.

**Lemma:** The four conditions in equation (2) in the text lead to the two relations

\[
\Omega(X_k - X_{k-1}, X_{k-1}, t_k) < \int_{X_{k-1}}^{X_k} \frac{\partial \Omega(0^+, x, t_k)}{\partial x} \, dx \quad \text{if } X_k < X_{k-1},
\]

\[
\Omega(X_k - X_{k-1}, X_{k-1}, t_k) < \int_{X_{k-1}}^{X_k} \frac{\partial \Omega(0^+, x, t_k)}{\partial x} \, dx \quad \text{if } X_k > X_{k-1},
\]

**Proof:** For \( X_k > X_{k-1} \), we have

\[
\Omega(X_k - X_{k-1}, X_{k-1}, t_k) = \int_{X_{k-1}}^{X_k} \frac{\partial \Omega(x - X_{k-1}, X_{k-1}, t_k)}{\partial x} \, dx < \int_{X_{k-1}}^{X_k} \frac{\partial \Omega(0^+, X_{k-1}, t_k)}{\partial x} \, dx \leq \int_{X_{k-1}}^{X_k} \frac{\partial \Omega(0^+, x, t_k)}{\partial x} \, dx.
\]

The strict inequality holds because marginal utility is strictly decreasing (assumption ii), and the second integrand is evaluated at \( 0^+ \), its smallest argument in the range of the integral. The weak inequality holds because marginal utility is weakly increasing in \( R \) (assumption iv) and each of the reference levels at which the integrand is evaluated are above \( X_{k-1} \). A similar proof holds when \( X_k < X_{k-1} \).

**Proof of Proposition 1:** For any given realized stochastic path of \( X \), assume there is a realized loss, and denote its reference level as \( R = X_{I-1} \). By assumption, \( X_I < X_{I-1} \). First, assume that the asset is eventually sold at a price above the reference level for this loss. Let \( J \) denote the first such sale; that is, for some \( J, X_J > X_{J-1} > X_{I-1} \). The total utility realized from the sales \( I \) through \( J \) is

\[
\sum_{k=I}^{J} \Omega(X_k - X_{k-1}, X_{k-1}, t_k) = \sum_{k=I}^{J-1} \Omega(X_k - X_{k-1}, X_{k-1}, t_k) + \Omega(X_{J-1} - X_{J-1}, X_{J-1}, t_J) + \Omega(X_J - X_{J-1}, X_{J-1}, t_J) - \Omega(X_{J-1} - X_{J-1}, X_{J-1}, t_J).
\]

The equality is a tautology with the utility from a fictitious sale occurring at time \( t_J \) at value \( X_{J-1} \) both added and subtracted. This lets us split the sales into two parts. The first line includes sales for which the gains and losses (though not their utility) net exactly to zero. The utility of this part is
\[
\sum_{k=I}^{J-1} \Omega(X_k - X_{k-1}, X_{k-1}, t_k) + \Omega(X_{J-1} - X_{J-1}, X_{J-1}, t_J)
\]

\[
< \sum_{k=I}^{J-1} \int_{x_{k-1}}^{x_k} \frac{\partial \Omega(0^+, x, t_k)}{\partial \Delta X} dx + \sum_{k=I}^{J-1} \int_{x_{k-1}}^{x_k} \frac{\partial \Omega(0^-, x, t_k)}{\partial \Delta X} dx + \int_{x_{J-1}}^{x_J} \frac{\partial \Omega(0^+, x, t_J)}{\partial \Delta X} dx < 0. \quad (A4)
\]

This contribution to utility is negative since each element of the first integrand over losses has a matching element in the second or third integrand of gains. Note that the integrals in the first sum are negative since the lower limits are larger than the upper limits. Also, as illustrated in Figure 1, the marginal utility of each loss element is greater than that of the matching gain element since marginal utility is decreasing in \(\Delta X\) (assumption ii) so evaluating the marginal utilities of gains at 0\(^+\) makes them bigger than the marginal utilities of losses evaluated at 0\(^-\), decreasing in time (assumption iii) and the gains occur later, and increasing in \(R\) (assumption iv) and the gains have smaller reference levels.

Since the terms in the first line of (A3) are negative, the total utility realized from the sales \(I\) through \(J\) is less than the sum of the two terms in the second line

\[
\sum_{k=I}^{J} \Omega(X_k - X_{k-1}, X_{k-1}, t_k) < \Omega(X_J - X_{J-1}, X_{J-1}, t_J) - \Omega(X_{J-1} - X_{J-1}, X_{J-1}, t_J)
\]

\[
= \int_{x_{J-1}}^{x_J} \frac{\partial \Omega(x - X_{J-1}, X_{J-1}, t_j)}{\partial \Delta X} dx < \int_{x_{J-1}}^{x_J} \frac{\partial \Omega(x - X_{J-1}, X_{J-1}, t_j)}{\partial \Delta X} dx \quad (A5)
\]

The new inequality in the second line holds because of assumptions (ii) and (iv). The final comparison shows that holding the stock with reference level \(X_{J-1}\) and selling at \(X_J\) lead to strictly higher utility than provided by the original plan, which, therefore, cannot be optimal.

Now suppose there is no later sale that occurs at a level above \(X_{J-1}\). In this case, the accumulated utility up through every sale \(J\) after \(I\) is negative. Precisely, using the same reasoning as before,

\[
\sum_{k=I}^{J} \Omega(X_k - X_{k-1}, X_{k-1}, t_k) < \sum_{k=I}^{J} \int_{x_{k-1}}^{x_k} \frac{\partial \Omega(0^+, x, t_k)}{\partial \Delta X} dx + \sum_{k=I}^{J} \int_{x_{k-1}}^{x_k} \frac{\partial \Omega(0^-, x, t_k)}{\partial \Delta X} dx < 0.
\]

This is true even for an infinite investment horizon with \(J \to \infty\). Therefore, holding the stock at a reference level of \(X_{J-1}\) with no subsequent sales is strictly better than the original plan.

To prove that, investors will voluntarily sell immediately to realize any gains in the absence of transactions costs, it is sufficient to show that any non-infinitesimal gain has more utility when split into two parts, i.e.,
\[ \Omega(X_t - X_{t-1}, X_{t-1}, t_f) < \Omega(X_s - X_{t-1}, X_{t-1}, t_s) + \Omega(X_s - X_s, X_s, t_f) \quad (A6) \]

Where \( X_s \) is any intermediate level between \( X_{t-1} \) and \( X_t \), and \( t_s \) is the time when the asset value reaches \( X_s \). Obtaining (A6) is immediate

\[
\begin{align*}
\Omega(X_t - X_{t-1}, X_{t-1}, t_f) &= \int_{X_{t-1}}^{X_t} \frac{\partial \Omega(x - X_{t-1}, X_{t-1}, t_f)}{\partial \Delta X} \, dx \\
&= \int_{X_{t-1}}^{X_s} \frac{\partial \Omega(x - X_{t-1}, X_{t-1}, t_f)}{\partial \Delta X} \, dx + \int_{X_s}^{X_t} \frac{\partial \Omega(x - X_{t-1}, X_{t-1}, t_f)}{\partial \Delta X} \, dx \\
&< \int_{X_{t-1}}^{X_s} \frac{\partial \Omega(x - X_{t-1}, X_{t-1}, t_s)}{\partial \Delta X} \, dx + \int_{X_s}^{X_t} \frac{\partial \Omega(x - X_s, X_s, t_f)}{\partial \Delta X} \, dx \\
&= \Omega(X_s - X_{t-1}, X_{t-1}, t_s) + \Omega(X_s - X_s, X_s, t_f)
\end{align*}
\]

The inequality holds because assumption (iii) together with \( t_s < t_f \) guarantee that the integrand in the first integral in the second line is less than that in the third line while assumptions (ii) and (iv) guarantee that the second integrand in the second line is less than that in the third line. The process of dividing the gain into smaller parts can continue ad infinitum so the conditions in (2) lead to immediate sales for all gains. \( \blacksquare \)

**B: Proof of Proposition 2**

As illustrated in Figure 4 there can be two local maxima to our optimization problem. The one-point maximum is a corner solution with \( \theta = 0 \); the two-point maximum is an interior maximum with \( \theta > 0 \). Which local maxima is the global maximum depends on the specific economic and utility parameters, but it can be most easily characterized by the loss aversion parameter, \( \lambda \). The value function for the one-point maximum does not depend on \( \lambda \) since no losses are ever realized. The value function for the two-point maximum is obviously decreasing in \( \lambda \). Therefore, there is a single critical value of \( \lambda \) at which the two value functions are equal that marks the change in regime.

Denote the two-point and one-point value functions as \( v^{(2)}(x) = C^{(2)}_1 x^{\gamma_1} + C^{(2)}_2 x^{\gamma_2} \) and \( v^{(1)}(x) = C^{(1)}_1 x^{\gamma_1} \). Since the value functions are equal everywhere that they are defined, we must have \( C^{(2)}_1 = C^{(1)}_1 \) and \( C^{(2)}_2 = 0 \). For convenience we repeat the boundary conditions and smooth-pasting conditions from (13) and (15) with \( C_2 = 0 \). They are

\[
\begin{align*}
C_1 \Theta^{\gamma_1} &= u(\kappa \Theta - 1) + (K \Theta)^\theta C_1 \\
\gamma_1 C_1 \Theta^{\gamma_1 - 1} &= \kappa u(\kappa \Theta - 1) + \beta K^\theta \Theta^{\theta - 1} C_1
\end{align*}
\]

and

\[
\begin{align*}
C_1 \Theta^{\gamma_1} &= u(\kappa \Theta - 1) + (K \Theta)^\theta C_1 \\
\gamma_1 C_1 \Theta^{\gamma_1 - 1} &= \kappa u(\kappa \Theta - 1) + \beta K^\theta \Theta^{\theta - 1} C_1
\end{align*}
\]

The two boundary conditions in (A8) give two formulae for \( C_1 \) which must give the same value
Similarly, the two equations for $\Theta$ in (A8) and (A9) and the two in $\theta$ also necessitate

\[ C_1 = \frac{u(\kappa \Theta - 1)}{\Theta^\gamma - (K \Theta)^\beta} = \frac{u(\kappa \theta - 1)}{\theta^\gamma - (K \theta)^\beta}. \]  

Substituting the utility function for scaled-TK utility into (A10) and solving for $\lambda$ gives

\[ \lambda = -\frac{(\kappa \Theta - 1)^{\alpha G} \theta^\gamma - K^\beta \Theta^\beta}{(1 - \kappa \theta)^{\alpha G} \Theta^\gamma - K^\beta \Theta^\beta}. \]  

Similarly the two equations in (A11) give

\[ 0 = (\alpha_G - \gamma_1)K^\beta \Theta^\gamma - (\alpha_G - \beta)K^\beta \kappa \Theta - \beta K^\beta \]
\[ 0 = (\alpha_L - \gamma_1)K^\beta \Theta^\gamma - (\alpha_L - \beta)K^\beta \kappa \theta - \beta K^\beta. \]  

These equations can be re-expressed as

\[ \Theta^\gamma - (K \Theta)^\beta = \frac{(K \Theta)^\beta (\gamma_1 - \beta)(\kappa \Theta - 1)}{(\alpha_G - \gamma_1)K \Theta + \gamma_1} \quad \theta^\gamma - (K \theta)^\beta = \frac{(K \theta)^\beta (\gamma_1 - \beta)(\kappa \theta - 1)}{(\alpha_L - \gamma_1)K \theta + \gamma_1}. \]  

Substituting back into (A12) gives

\[ \lambda = \frac{(\kappa \Theta - 1)^{\alpha G} - 1 \theta^\gamma - K^\beta \theta^\beta}{(1 - \kappa \Theta)^{\alpha G} - 1 \Theta^\gamma - K^\beta \Theta^\beta}, \]  

which is the desired expression in (16).

Correspondingly for modified-TK utility we have from (A10) instead of (A12)

\[ \lambda = \frac{\alpha_L (\kappa \Theta)^{\alpha G} - 1 \Theta^\gamma - K^\beta \Theta^\beta}{\alpha_G (\kappa \Theta)^{\alpha G} - 1 \Theta^\gamma - K^\beta \Theta^\beta}, \]  

And from (A11) in place of (A13)

\[ 0 = (\alpha - \gamma_1)K^\beta \Theta^\gamma - (\alpha_G - \beta)K^\beta (\kappa \Theta)^{\alpha G} - \beta K^\beta \]
\[ 0 = (\alpha_L - \gamma_1)K^\beta \Theta^\gamma - (\alpha_L - \beta)K^\beta (\kappa \theta)^{\alpha L} - \beta K^\beta. \]  

from which we can derive
\[ \lambda = \frac{\alpha_L}{\alpha_G} \left( \frac{\theta}{\Theta} \right) \left( \frac{(\alpha_G - \gamma_1)\kappa^{\alpha_G} \Theta^{\alpha_G} + \gamma_1}{(\alpha_L - \gamma_1)\kappa^{\alpha_L} \Theta^{\alpha_L} + \gamma_1} \right) \] (A18)

which is the desired expression in (19).

C: Verification of the Optimality of Constant Proportional Sales Policies

Consider the general sale and reinvestment problem given in the text. For any given realized stochastic price path and sales policy denote the original stock price and the stock prices at the points of sale as \( S_0, S_1, \ldots, S_n, \ldots \). The dollar amount under investment is denoted as \( X_n^{\pm} \) where the minus and positive superscripts denote the value before and after paying the transactions costs at the \( n \)th sale and repurchase. \( N_n \) denotes the number of shares purchased in the \( n-1 \)st purchase and sold in the \( n \)th sale. The relations among the \( X \)'s are given by the recurrences

\[ X_n^+ = (1 - k_s)X_n^-/(1 + k_p) \equiv KX_n^- \quad X^+_{n+1} = X_n^+S_{n+1}/S_n. \] (A19)

It is readily verified that

\[ X_n^{+j} = N_nS_n(S_{n+j}/S_n)\ldots(S_{n+j}/S_{n+j-1})K^{j-1}/(1 + k_p) = N_nS_{n+j}K^{j-1}/(1 + k_p) \]

\[ X_n^{+j} = KX_n^{-j} = N_nS_{n+j}K^j/(1 + k_p). \] (A20)

The reference level, \( R_n \), established after the \( n \)th sale and repurchase is proportional to both \( X_n^{\pm} \). We have assumed that the investor considers both sets of transactions costs in setting his reference level with \( R_n = X_n^+ = KX_n^- \). Alternatively he might ignore the repurchase costs setting \( R_n = X_n^-/(1 - k_s) = X_n^+(1 + k_p). \) To cover both of these and many other cases we define \( R_n = K'X_n^+ \).

The subjective realized rate of return gain at the \( n \)th sale is

\[ \kappa X_n^-/R_n-1 = (\kappa/K')X_n^-/X_n^+ - 1 = (\kappa/K')S_n/S_n-1 - 1. \] (A21)

The utility burst at time \( t_n \) is

\[ e^{-\delta t_n}R_n^\beta u(\kappa X_n^-/R_n-1) \] (A22)

The argument of the utility function and the reference level are

\[ \kappa X_n^- / R_n - 1 = (\kappa/K')S_n/S_n - 1 \]

\[ R_n = K'X_n^+ = \frac{K'N_0S_{n-1}K^{n-1}}{1 + k_p} = K'X_0^nK^{n-1}S_{n-1}/S_0 = K^{n-1}R_0S_{n-1}/S_0. \] (A23)

So for this particular realized outcome and strategy, lifetime future utility measured just after the \( n \)th sale and repurchase is
Now consider a rule that generates the optimal sales policy. It sets gain and loss sales points for the first sale of $S_1(S_0)$ and $S_1(S_0)$ and contingent rules for the second sales points, $S_2(S_t)$ and $S_2(S_t)$, etc. These optimal policies cannot depend on time since the stochastic process is time homogeneous and the investor has a constant subjective discount rate. Nor can the optimal policies depend on the reference level, $R_{nt}$, (separately from the prevailing stock price, $S_n$) since looking ahead from time $t_n$, the reference level affects the realized utility only through the proportional scaling factor $\beta$. Finally, the optimal policies for the second sale can depend on $S_0$ only through $S_1$ since the former price doesn’t affect the utility looking forward from time $t_1$ and the stock returns are independent and identically distributed. Therefore we must have $S_2(S_t) / S_1 = S_1(S_0) / S_0$, etc.

It is clear from this analysis that constant proportional transactions costs, a constant and proportional subjective interpretation of the realized return and reference level ($\kappa$ and $K'$), a stock price process with independent and identically distributed returns, a constant rate of time preference, an infinite horizon, and a utility based on rates of return with power scaling factor are all necessary for a constant optimal policy.

**D: Transversality Conditions**

There are several different transversality conditions required in our analysis to keep the value function finite and produce a well-defined optimal trading strategy. First the discount rate must be large enough and the growth rate of the asset must not be too large or utility bursts far in the future dominate the value function and make its unbounded. These two conditions mimic the transversality conditions in the standard portfolio problem. Second the scaling parameter $\beta$ must be not too large otherwise repeated selling lets total utility accumulate too quickly by increasing the reference level and thereby the utility bursts of future sales. Finally, for scaled-TK utility, transactions costs cannot be zero or the investor can repeatedly realize small gains with their very high marginal utility.

A simple version of the limit on $\beta$ was seen in equation (21). Here we examine the general case when the reference level grows at the rate $\zeta$ between sales and there may be an exogenous Poisson event which terminates participation. Assuming a constant sales policy, $\Theta$, the reference level at the time of the $n$th sale which was established at the time of the $(n-1)$th sale will be $R(t_n) = R_0 e^{\zeta t_n} (K \Theta)^{n-1}$. The $n$th sale has a subjective gain of $(\kappa \Theta - 1) R(t_n)$ or $\kappa \Theta - 1$ per dollar of the reference level. The probability that trading has not been terminated before the $n$th sale is $e^{-\lambda t_n}$. The expected lifetime utility from a series of sales at gains and no sales at losses is

$$
\mathbb{E} \sum_{n=1}^{\infty} e^{-(p+\delta)\tilde{t}_n} R^n(\tilde{t}_n) u(\kappa \Theta - 1) = u(\kappa \Theta - 1) \frac{R_0^\beta}{(K \Theta)^n} \sum_{n=1}^{\infty} (K \Theta)^{n \beta} \mathbb{E} e^{(\kappa \Theta - 1 - \delta)\tilde{t}_n}. \quad (A25)
$$

Here $\tilde{t}_n$ is the random time of the $n$th sale. Note that the utility bursts realized are known; it is only the timing that is random. Since the times between successive sales are independent and identically distributed, the above expression can be simplified to

$$
\mathbb{E} \sum_{n=1}^{\infty} e^{-(p+\delta)\tilde{t}_n} R^n(\tilde{t}_n) u(\kappa \Theta - 1) = u(\kappa \Theta - 1) \frac{R_0^\beta}{(K \Theta)^n} \sum_{n=1}^{\infty} (K \Theta)^{n \beta} \mathbb{E} e^{(\kappa \Theta - 1 - \delta)\tilde{t}_n}.
$$
identically distributed

\[
\mathbb{E}[e^{(\beta\zeta - \delta)\hat{t}_n}] = \mathbb{E}[e^{(\beta\zeta - \delta)\hat{t}_{n-1}}] \cdots \mathbb{E}[e^{(\beta\zeta - \delta)(\hat{t}_2 - \hat{t}_1)}] = \left(\mathbb{E}[e^{(\beta\zeta - \delta)\hat{t}_1}]\right)^n, \tag{A26}
\]

and the final sum in (A25) is

\[
\sum_{n=1}^{\infty} \left( (K\Theta)^{\beta_n} \mathbb{E}[e^{(\beta\zeta - \delta)\hat{t}_1}] \right)^n \tag{A27}
\]

which converges if and only if \((K\Theta)^{\beta} \mathbb{E}[e^{(\beta\zeta - \delta)\hat{t}_1}] < 1\). The expected value in (A27) depends on \(\Theta\) through the stopping time \(\hat{t}_1\), but unless \(\delta + \rho > \beta\zeta\), it is at least one for any \(\Theta\) so the sum is unbounded for any \(\Theta > 1/K\) and infinite utility is possible. This condition corresponds to the standard transversality condition that the discount rate must be positive.

The expectation in (A27) is a particular value of the Laplace transform\(^45\) of the first-passage time density of the random variable, \(X_t/R_t\), to \(\Theta\); that is,

\[
\mathbb{E}[e^{-\hat{t}_1}] = \exp\left(\frac{\ell n\Theta}{\sigma^2} \left[ \mu' - \sqrt{(\mu')^2 + 2\ell \sigma^2} \right] \right) \quad \text{where} \quad \mu' = \mu - \frac{1}{2} \sigma^2
\]

\[
\Rightarrow \quad \mathbb{E}[e^{(\beta\zeta - \delta)\hat{t}_1}] = \exp(-\gamma_1\ell n\Theta) \tag{A28}
\]

where \(\gamma_1\) is the positive exponent in the solution of the differential equation for valuation as given in (37). Therefore, when \(\delta + \rho > \beta\zeta\) we require

\[
1 > (K\Theta)^{\beta} \mathbb{E}[e^{(\beta\zeta - \delta)\hat{t}_1}] = (K\Theta)^{\beta} \Theta^{-\gamma_1}. \tag{A29}
\]

If \(\beta > \gamma_1\), this inequality is violated for any \(\Theta > K^{-\beta/(\beta-\gamma_1)}\) so there is no well-defined optimum. If \(0 \leq \beta \leq \gamma_1\), then the sum converges for all feasible \(\Theta\) (i.e., \(\Theta > 1/\kappa \geq 1\)).\(^46\) Combining these results, necessary conditions for there to be no investment plans which lead to infinite expected utility are

\[
0 \leq \beta < \gamma_1 \quad \text{and} \quad \beta\zeta < \delta + \rho. \tag{A30}
\]

However, these conditions are not sufficient as unbounded utility might also be achieved by waiting for a very long time before selling once if the growth rate of the asset is large. Consider the model as described in the text. The reference level is growing at the rate \(\zeta\), utility burst of a sale is \(R_t^{\beta}(\kappa X_t/R_t - 1)\), and the investment value is evolving as \(dX_t/X_t = \mu dt + \sigma d\omega\). Define \(\zeta\)

\(^{45}\) The Laplace transform of a density function, \(\mathcal{L}(f(t)) \equiv \mathbb{E}[e^{-t}]\), can be easily determined from the moment generating function, \(\mathcal{M}(f(t)) \equiv \mathbb{E}[e^{t}]\), for a negative argument. The Laplace transform is defined for all values of \(\mu\) (unlike the moment generating function) since the first passage time is a positive random variable.

\(^{46}\) We have assumed that \(\beta \geq 0\) to ensure participation; however, if \(\beta < 0\), then the sum might diverge for \(\Theta = 1\). The minimum value for \(\Theta\) is \(1/\kappa\), so the convergence condition is \(1 > K^{-\beta}(\kappa^{-\beta})\). If the investor is fully cognizant of all transactions costs then \(\kappa = K = (1 - k_p)/(1 + k_p)\), and this condition is met; otherwise it may not be. For example, if the investor compares his net gain after sales costs but before reinvestment costs \((\kappa = 1 - k_p)\), the convergence condition is \((1 + k_p)^{\beta} > (1 - k_p)^{\beta}\).
\[ Y_t = \int_0^\infty e^{-(\mu + \delta)\xi_t + R_t^2 (\xi_t - \frac{1}{2} \sigma^2 t)} d\xi_t, \] 

where \( f \) is the probability density function of \( \xi_t \). From our previous analysis, we need only consider the case when \( \beta \xi < \delta + \rho \). With this restriction, expected utility from this single gain can be unbounded only if the utility function itself is unbounded. That is, for modified-TK utility, we would require \( \alpha_G > 0 \) to have a transversality violation.

For both utility functions \( u(\exp(\xi_t) - 1) \sim \exp(\alpha_G \xi_t) \) for large \( \xi_t \), which is normally distributed with mean \( (\mu - \zeta - \frac{1}{2} \sigma^2) t \) and variance \( \sigma^2 t \), so

\[ Y_t = \frac{R_t^b}{\sqrt{2\pi\sigma^2 t}} \int_0^\infty \exp\left( (\beta \xi - \delta - \rho + \alpha_G [\mu - \zeta + (\alpha_G - 1) \frac{1}{2} \sigma^2]) t \right) \frac{\xi_t - (\mu - \zeta - \frac{1}{2} \sigma^2) t}{2\sigma^2 t} d\xi_t \]

\[ = \exp(\beta \xi - \delta - \rho + \alpha_G [\mu - \zeta + (\alpha_G - 1) \frac{1}{2} \sigma^2]) t \]

\[ \times \frac{R_t^b}{\sqrt{\sigma^2 t}} \int_0^\infty \phi\left( \frac{\xi_t - (\mu - \zeta + (\alpha_G - 1) \frac{1}{2} \sigma^2) t}{\sigma \sqrt{t}} \right) d\xi_t \]

\[ = R_t^b \exp(\beta \xi - \delta - \rho + \alpha_G [\mu - \zeta + (\alpha_G - 1) \frac{1}{2} \sigma^2]) t \Phi\left( \frac{(\mu - \zeta + (\alpha_G - 1) \frac{1}{2} \sigma^2) \sqrt{t}}{\sigma} \right). \]

If the exponent in the first term is negative,

\[ \alpha_G [\mu - \zeta + \frac{1}{2} (\alpha_G - 1) \sigma^2] < \delta + \rho - \beta \xi. \]  

then the limit of the first term is zero, and utility must be bounded since the cumulative normal is bounded. If this exponent is positive, utility might still be bounded if the numerator in the cumulative normal is negative so its argument goes to \(-\infty\) and its limit is 0. But if the numerator is negative and \( \alpha_G > 0 \), then

\[ 0 > \mu - \zeta + (\alpha_G - 1) \sigma^2 > \mu - \zeta + \frac{1}{2} (\alpha_G - 1) \sigma^2 \]

\[ \text{(A34)} \]

and equation (A33) must be satisfied since \( \delta + \rho - \beta \xi \) is positive by (A30), so a negative numerator does not add any additional violations.

Therefore, sufficient conditions for no transversality violation of this type are either

\[ \alpha_G < 0 \quad \text{or} \quad \alpha_G [\mu - \zeta + \frac{1}{2} (\alpha_G - 1) \sigma^2] < \delta + \rho - \beta \xi. \]  

\[ \text{(A35)} \]

\[ ^{47} \text{A similar proof applies when } \alpha_G = 0 \text{ which corresponds to log utility of gains. Since } \gamma_i \text{ is positive the condition in (A36) is satisfied.} \]
This can be simplified and expressed as

\[ \alpha_G < \gamma_1. \]  

(A36)

Note that assuming \( \alpha_G < 1 \) expands the set of admissible values for \( \mu \). For risk neutrality, we must have \( \mu < \delta + \zeta (1-\beta) \).

Finally unbounded utility can be achieved for scaled-TK utility with \( \alpha_G < 1 \) if there are no transactions cost and no Poisson event that ends trading prematurely. We construct a sequence of sales strategies as the transaction costs go to zero, and we demonstrate that the constructed strategies lead to unbounded utility. As a result, the maximizing strategy must yield infinite utility in the limit of zero costs.

Consider a sequence of economies indexed by \( k \). We assume that the sales and purchase costs are in the same proportion \( c \geq 0 \) for each step of the sequence, \( k_s = k \) and \( k_p = ck \). At each step the not necessarily optimal sales policies are

\[ \Theta = 1 / (1-k) + k^{\omega_0}, \quad \theta = 1 - k^{\omega_0}, \quad \text{where} \quad 0 < \omega_0, \omega_1 < \frac{1}{2}. \]  

(A37)

Initial utility is \( \nu(1) = C_1 + C_2 \) where these constants are defined in (14). Using a Taylor expansion for small \( k \), the functions determining \( C_1 \) and \( C_2 \) are

\[
\begin{align*}
    c_1(\Theta) &= (\gamma_1 - \beta)k^{\omega_0} + \frac{1}{2}[\gamma_1(\gamma_1 - 1) - \beta(\beta - 1)]k^{2\omega_0} + o(k^{2\omega_0}) \\
    c_1(\theta) &= (\beta - \gamma_1)k^{\omega_0} + \frac{1}{2}[\gamma_1(\gamma_1 - 1) - \beta(\beta - 1)]k^{2\omega_0} + o(k^{2\omega_0}) \\
    c_2(\Theta) &= (\gamma_2 - \beta)k^{\omega_0} + \frac{1}{2}[\gamma_2(\gamma_2 - 1) - \beta(\beta - 1)]k^{2\omega_0} + o(k^{2\omega_0}) \\
    c_2(\theta) &= (\beta - \gamma_2)k^{\omega_0} + \frac{1}{2}[\gamma_2(\gamma_2 - 1) - \beta(\beta - 1)]k^{2\omega_0} + o(k^{2\omega_0}).
\end{align*}
\]

(A38)

The constants are

\[
\begin{align*}
    C_1 &= \frac{(\beta - \gamma_2)k^{\omega_0 + \omega_1\omega_0} - \lambda(\beta - \gamma_2)k^{\omega_0 + \omega_0\omega_0} + o(k^{\min(\omega_0 + \omega_0, \omega_0 + \omega_0)})}{C(k^{2\omega_0 + \omega_0} + k^{2\omega_0 + \omega_0}) + o(k^{\min(2\omega_0 + \omega_0, 2\omega_0 + \omega_0)})} \\
    C_2 &= \frac{(\gamma_1 - \beta)k^{\omega_0 + \omega_1\omega_0} - \lambda(\gamma_1 - \beta)k^{\omega_0 + \omega_0\omega_0} + o(k^{\min(\omega_0 + \omega_0, \omega_0 + \omega_0)})}{C(k^{2\omega_0 + \omega_0} + k^{2\omega_0 + \omega_0}) + o(k^{\min(2\omega_0 + \omega_0, 2\omega_0 + \omega_0)})}
\end{align*}
\]

(A39)

where

\[
\begin{align*}
    C &= \frac{1}{2}(\gamma_2 - \gamma_1)[\gamma_1\gamma_2 - \frac{1}{2}\beta(\gamma_1 + \gamma_2) + \frac{1}{2}\beta^2] \\
    &\geq \frac{1}{2}(\gamma_2 - \gamma_1)(\gamma_1\gamma_2 - \frac{1}{2}\beta\gamma_1 - \beta\gamma_2 + \frac{1}{2}\beta^2) = \frac{1}{2}(\gamma_1 - \gamma_2)(1 - \beta)(\frac{1}{2}\beta - \gamma_2) > 0.
\end{align*}
\]

If we choose the converging rates, \( \omega_0 \) and \( \omega_1 \), such that \( \omega_0/\omega_1 > (1-\alpha_L)/(1-\alpha_G) \), then from (A39) it is easy to verify that \( C_1, C_2 \to \infty \), as \( k \to 0 \). The truly optimal value function \( \nu \), must satisfy \( \nu(1) \)

\[ \text{Given that} \quad \omega_0 \text{ and } \omega_1 \text{ are both less than } \frac{1}{2}, \text{ all the terms associated with } c \text{ are included in the higher-order terms } o(k^{2\omega_0}) \text{ and } o(k^{2\omega_0}). \]
\[ \geq C_1 + C_2, \text{ and therefore must goes to } \infty \text{ in the limit.} \]

**E: Solving the Differential Equations for Exponentially Smoothed Reference Point**

The partial differential equation for an exponentially-smoothed reference point is given in equation (27) as

\[ 0 = \frac{1}{2} \sigma^2 X^2 Y_{xx} + \mu X Y_X + \eta (X - R) Y_R + Y_t. \]  
(A40)

The homogeneity property still holds so we have \( Y(X, R, t) = e^{-\beta t} R^\beta \nu(x) \), giving a reduced-form differential equation of

\[ 0 = \frac{1}{2} \sigma^2 x^2 \nu'(x) + (\mu + \eta - \eta x) x \nu'(x) + (\eta x - \delta - \beta \eta) \nu(x). \]  
(A41)

Making the change of variable \( \nu(x) = x^{\eta/2} g(z) \) where \( z \equiv 2 \eta x / \sigma^2 \), \( q \equiv [1 - 2(\mu + \eta) / \sigma^2] \pm h \), and \( h \equiv \sigma^2 \sqrt{(2 \mu + 2 \eta - \sigma^2)^2 + 8 \sigma^2 (\delta + \beta \eta)} \), then \( g(z) \) is the solution to

\[ 0 = zg''(z) + (1 \pm h - z) g'(z) + (1 - 2q) g(z). \]  
(A42)

This will be recognized as the confluent hypergeometric equation of Kummer with solution

\[ D_1 M (2q - 1, 1 \pm h, z) + D_2 z^{\frac{h}{2}} M (2q - 1 \mp h, 1 \mp h, z) \]

where \( M(a, b, z) = \frac{\Gamma(b)}{\Gamma(a)} \sum_{n=0}^{\infty} \frac{(a + n) z^n}{\Gamma(b + n) n!} \)  
(A43)

and \( \Gamma(a) \) is the gamma or generalized factorial function. In terms of the original variable \( x \), the solution is

\[ C_1 x^{\eta/2} M (2q - 1, 1 \pm h, z) + C_2 x^{\eta/2} z^{\frac{h}{2}} M (2q - 1 \mp h, 1 \mp h, z) \]

where \( z \equiv \frac{2 \eta x}{\sigma^2} \), \( q \equiv \frac{1 - \frac{2(\mu + \eta)}{\sigma^2}}{\sigma^2} \pm h \), and \( h \equiv \frac{\sqrt{(2 \mu + 2 \eta - \sigma^2)^2 + 8 \sigma^2 (\delta + \beta \eta)}}{\sigma^2} \)  
(A44)

which is (28) in the text.

**F: Solving the Inhomogeneous Differential Equations for Ongoing Utility**

The solution of the homogeneous differential equation (10) is standard. Here we derive the solutions of inhomogeneous equations (30) and (38).\(^{49}\) For modified-TK utility, the inhomogeneous terms are \([ (\tilde{\kappa} x)^{\alpha} - 1 ] / \alpha \tilde{G} \) and \(- \lambda [1 - (\tilde{\kappa} x)^{\alpha\ell} ] / \alpha \ell \) in the gain and loss regions, respectively. Both have the form \( a + bx^\alpha \). Using the method of undetermined coefficients, we can express a particular solution to the equation as \( A + Bx^\alpha \). Substituting into (38) gives

\(^{49}\) The two equations differ only in the two \( \rho \)'s. For ongoing utility \( \rho = 0 \). For Poisson events, \( \dot{\rho} = \rho \).
\[ 0 = \frac{1}{2} \sigma^2 x^\alpha B \alpha (\alpha - 1) + (\mu - \zeta) x^\alpha B \alpha - (\rho + \delta - \zeta \beta)(A + B x^\alpha) + \hat{\rho}(a + bx^\alpha). \] (A45)

Since this must be zero for all \( x \), the constant and \( x^\alpha \) terms must separately be zero, and we can solve for \( A \) and \( B \) as\(^{50}\)

\[
A = \frac{-\hat{\rho} a}{\rho + \delta - \zeta \beta} \quad \quad B = \frac{-\hat{\rho} b}{\frac{1}{2} \sigma^2 \alpha (\alpha - 1) + (\mu - \zeta) \alpha - (\rho + \delta - \zeta \beta)}. \] (A46)

For scaled-TK utility, the inhomogeneous terms are \((\hat{k} x - 1)^{a_G}\) and \(-\lambda(1 - \hat{k} x)^{a_\ell}\) in the gain and loss regions. We employ the Wronskian method. The two homogeneous solutions are \( x^{\gamma_1} \) and \( x^{\gamma_2} \). The determinant of the Wronskian is \( \gamma_2 x^{\gamma_1} x^{\gamma_2} - \gamma_1 x^{\gamma_1} x^{\gamma_2} = (\gamma_2 - \gamma_1) x^{\gamma_1 + \gamma_2 - 1} \equiv \mathcal{W}(x) \). A particular solution in the loss region is

\[
v_L(x) = Q_L(x) x^{\gamma_1} + P_L(x) x^{\gamma_2}
\]

where

\[
Q_L(x) = \int_{1/\hat{k}}^{x} \frac{-2z^{\gamma_1} \hat{\rho}\gamma(1 - \hat{k} z)^{a_G}}{\sigma^2 z^2 \mathcal{W}(z)} \, dz = \frac{-2\hat{\rho} \gamma}{\sigma^2 (\gamma - \gamma_1)} \int_{1/\hat{k}}^{x} z^{\gamma_1 - 1} (\hat{k} z - 1)^{a_G} \, dz
\]

\[
= \frac{2\hat{\rho} \gamma \hat{k}^{\gamma_1}}{\sigma^2 (\gamma - \gamma_1)} \int_{1/\hat{k}}^{x} \Gamma_{\gamma_1 - a_G - 1} (1 - t)^{a_G} \, dt = \frac{2\hat{\rho} \gamma \hat{k}^{\gamma_1}}{\sigma^2 (\gamma - \gamma_1)} B((\hat{k} x)^{-1}; \gamma_1 - a_G, \alpha_G + 1) + \text{const}
\]

\[
P_L(x) = \int_{1/\hat{k}}^{x} \frac{2z^{\gamma_2} \hat{\rho}(1 - \hat{k} z)^{a_\ell}}{\sigma^2 z^2 \mathcal{W}(z)} \, dz = \frac{2\hat{\rho} \gamma}{\sigma^2 (\gamma - \gamma_1)} \int_{1/\hat{k}}^{x} z^{\gamma_2 - 1} (\hat{k} z - 1)^{a_\ell} \, dz
\]

\[
= \frac{-2\hat{\rho} \gamma \hat{k}^{\gamma_2}}{\sigma^2 (\gamma - \gamma_1)} B((\hat{k} x)^{-1}; \gamma_2 - a_G, \alpha_G + 1) + \text{const}.
\]

\(^{50}\) Note that the denominator for \( B \) is the characteristic equation for the two roots, \( \gamma_1 \) and \( \gamma_2 \), so when \( \alpha \) is equal to either \( \gamma_1 \) or \( \gamma_2 \), this particular solution is not valid. In this case, the particular solution is \( A + B x^\alpha \ln(x) \) with \( A \) the same and \( B \equiv -\hat{\rho} b/[(\mu - \zeta + \sigma^2 \alpha (\alpha - \frac{1}{2}))]. \)

\(^{51}\) See Abramowitz and Stegun (1970) for this definition and other details about the incomplete beta function.
Since $Q_G$ and $P_G$ are multiplied by $x^m$ and $x^n$ which are the homogeneous solutions, the constants can be absorbed into the constants of the homogeneous solution.
References


Figure 1: Illustration of Proposition 1 — No Voluntary Sales at Losses. This figure illustrates the proof of Proposition 1. Each portion of every loss is matched against a portion from later gain(s). The corresponding loss and gain portions are of equal size, but the gain portion has smaller marginal utility throughout due to risk aversion, occurs later, and has a lower reference level. Each of these properties makes the total utility from the gains less than the disutility of the losses.
Figure 2: Determination of the Scaling Parameter $\beta$. This figure illustrates the determination of the scaling parameter for a utility function $R^\beta u(\Delta x)$. The utility function is defined for the change in value per dollar of the reference level, $R$, $\Delta x = \Delta X/R$. The two illustrated utility functions are scaled-Tversky-Kahnemann and modified-Tversky-Kahnemann as defined in equations (5) and (6). The x-axis is the value $z$ that answers the question: “Gaining $10 relative to a reference level of $100 makes just as happy as gaining $z$ relative to a reference level of $1000.” The parameter $\alpha$ is $\alpha_G$ if the question is phrased about gains and $\alpha_L$ if phrased about losses.
Figure 3: Determination of the Optimal Sales Policies. This figure illustrates the value function and the optimal policy for realizing gains and losses. The value in the continuation region is tangent to the payoff functions at both sales points, $\theta$ and $\Theta$. In this illustrated case, it is optimal to realize some losses.
Figure 4: The Value Function for Scaled-TK Utility. The reduced value function for scaled-TK utility measured immediately after a sale and reference level reset, $v(1)$, is plotted against different loss sales points, $\theta$. The gain sales point, $\Theta$, is fixed at its optimal value. The solid black line shows the value function for $\lambda = 2.892$, the dashed red line shows the value function for $\lambda = 2.8$, and the dotted blue line shows the value function for $\lambda = 3.0$. The other parameters are $\mu = 10\%$, $\sigma = 30\%$, $k_s = k_p = 1\%$, $\alpha_G = \alpha_L = 0.5$, $\beta = 0.25$, $\zeta = 5\%$, $\delta = 4\%$ $\kappa = 1-k_s$. For $\lambda = 2.8$, the two-point policy ($\theta = 0.19$, $\Theta = 1.03$) is optimal. For $\lambda = 3$, the one-point policy ($\Theta = 1.03$) is optimal. For the critical value $\lambda = 2.892$, the two point policy ($\theta = 0.16$, $\Theta = 1.03$) and the one-point policy ($\Theta = 1.03$) have identical utility. (The optimal $\Theta$ are not identical, but differ in the 4th or 5th decimal place.)
Figure 5: The Value Function. The initial optimized value function for scaled-TK and modified-TK utility plotted against $\mu$ and $\sigma$. The default parameters are $\mu = 10\%$, $\sigma = 30\%$, $k_s = k_p = 1\%$, $\beta = 0.25$, $\lambda = 2$, $\delta = 4\%$. For scaled-TK utility $\alpha_G = \alpha_L = 0.5$. For modified-TK utility $\alpha_G = -3$, $\alpha_L = 5$. The investor ignores the reinvestment cost in assessing his realized gains ($\kappa = 1-k_s$).
Figure 6: Indifference Curves. The indifference curves for scaled-TK and modified-TK utility. The panels on the left use the default parameters, $\beta = 0.25$, $\lambda = 2$, $\zeta = 5\%$, $\delta = 4\%$, $k_s = k_p = 1\%$, and $\alpha_G = \alpha_L = 0.5$ for scaled-TK utility and $\alpha_G = -3$, $\alpha_L = 5$ for modified-TK utility. The panels on the right use $\beta = 0$, $\zeta = 0$, $\delta = 3\%$, and $\alpha_G = -1$, $\alpha_L = 3$ for modified-TK utility. The investor ignores the reinvestment cost in assessing his realized gains ($\kappa = 1-k_s$).
Figure 7: The Optimal Sales Policies for Scaled-TK Utility. The optimal sales points, Θ and θ, for scaled-TK utility plotted against various parameters. The default parameters in each graph are μ = 10%, σ = 30%, k_s = k_p = 1%, α_G = α_L = 0.5, β = 0.25, λ = 2, ζ = 5%, δ = 4%. The dotted blue lines show the optimal policies for an investor who ignores the reinvestment cost in assessing his gains (κ = 1−k_s). The solid red lines show the optimal policies for an investor who recognizes the reinvestment cost in assessing his gains κ = K = (1−k_s)/(1+k_p).
Figure 8: The Optimal Sales Policies for Modified-TK Utility. The optimal sales points, $\Theta$ and $\theta$, for modified-TK utility plotted against various parameters. The default parameters in each graph are $\mu = 10\%$, $\sigma = 30\%$, $k_s = k_p = 1\%$, $\alpha_G = -3$, $\alpha_L = 5$, $\beta = 0.25$, $\lambda = 2$, $\zeta = 5\%$, $\delta = 4\%$. The dotted blue lines show the optimal policies for an investor who ignores the reinvestment cost in assessing his gains ($\kappa = 1 - k_s$). The solid red lines show the optimal policies for an investor who recognizes the reinvestment cost in assessing his gains, $\kappa = K = (1-k_s)/(1+k_p)$.
Figure 9: The Optimal Sales Points and Value Function for an Exponentially Smoothed Reference Level. The optimal sales points, $\Theta$ and $\theta$, and the value function, $v(1)$, for scaled-TK utility plotted against $\eta$, the smoothing parameter. The investor ignores the reinvestment cost in assessing his realized gains ($\kappa = 1-k_s$). The default parameters, $\sigma = 30\%$, $k_s = k_p = 1\%$, $\alpha_G = \alpha_L = 0.5$, $\beta = 0.25$, $\lambda = 2$ are used. The reference level growth rate $\zeta$ is not defined for this process so the parameters $\mu$ and $\delta$ are adjusted to their effective levels, $\mu - \zeta = 5\% \rightarrow \mu$, $\delta - \zeta \beta = 2.75\% \rightarrow \delta$ are used. The solid red lines show the sales points $\Theta$ and $\theta$. The dotted green line shows the value function, $v(1)$. 
Figure 10: The Optimal Sales Points for Realization Utility with Ongoing Utility for Paper Gains and Losses. The optimal sales points, $\Theta$ and $\theta$ (heavy lines), and value function, $v(1)$ (light lines), for modified-TK utility plotted against $\hat{\rho}$, the parameter measuring the importance of ongoing utility relative to the utility bursts of sales. The parameters are the default values $\mu = 10\%$, $\sigma = 30\%$, $k_s = k_p = 1\%$, $\beta = 0.25$, $\lambda = 2$, $\zeta = 5\%$, $\delta = 4\%$. The investor ignores the reinvestment cost in assessing his realized gains ($\kappa = 1 - k_s$) and ignores transactions costs in his subjective view of ongoing gain or loss ($\hat{\kappa} = 1$). The blue lines show the local maximum with $\Theta$ very high. The red lines show the local maximum with $\Theta \approx 1$. The dotted portions of each line show indicate the local maximum is not a global maximum.
Figure 11: The Participation Constraint for Realization and Ongoing Utility for Paper Gains and Losses. Each curve shows the minimum $\mu$ at which the investor is willing to participate as a function of $\sigma$. The parameter, $\hat{\rho}$, measures the importance of ongoing utility relative to the utility bursts of sales. The investor ignores transactions costs in his subjective view of ongoing gain or loss, $\hat{\kappa} = 1$. The other parameters are $\mu = 10\%$, $\sigma = 30\%$, $k_s = k_p = 1\%$, $\alpha_G = -3$, $\alpha_L = 5$, $\beta = 0.25$, $\lambda = 2$, $\zeta = 5\%$, $\delta = 4\%$. The investor ignores the reinvestment cost in assessing his realized gains ($\kappa = 1 - k_s$).
Figure 12: The Participation Constraint with Poisson Exit Events. Each curve shows the minimum $\mu$ at which the investor is willing to participate as a function of $\sigma$. The parameter, $\rho$, is the Poisson intensity of forced sales. The other parameters are $\mu = 10\%$, $\sigma = 30\%$, $k_s = k_p = 1\%$, $\alpha_G = -3$, $\alpha_L = 5$, $\beta = 0.25$, $\lambda = 2$, $\zeta = 5\%$, $\delta = 4\%$. The investor ignores the reinvestment cost in assessing his realized gains ($\kappa = 1 - k_s$).
Figure 13: The Trading Probabilities. The trading probabilities for modified-TK utility plotted against $\mu$ and $\sigma$. $Q_\theta$ and $Q_\Theta$ are the probabilities of trading at a loss and a gain respectively within a year of the original purchase. The default parameters are $\mu = 10\%$, $\sigma = 50\%$, $k_s = k_p = 1\%$, $\alpha_G = -3$, $\alpha_L = 8$, $\beta = 0.25$, $\lambda = 1.5$, $\zeta = 5\%$, $\delta = 4\%$, $\rho = 0$. The solid lines show the effect of $\mu$ and $\sigma$ on trading probabilities at gains and losses with endogenously changing sales points, $\theta$ and $\Theta$. The dotted lines show the effect of $\mu$ and $\sigma$ on trading probabilities but with $\theta$ and $\Theta$ fixed at their default values.