Realization Utility

Nicholas Barberis and Wei Xiong
Yale University and Princeton University

October 2011*

Abstract

A number of authors have suggested that investors derive utility from *realizing* gains and losses on assets that they own. We present a model of this "realization utility," analyze its predictions, and show that it can shed light on a number of puzzling facts. These include the disposition effect, the poor trading performance of individual investors, the higher volume of trade in rising markets, the effect of historical highs on the propensity to sell, the individual investor preference for volatile stocks, the low average return of volatile stocks, and the heavy trading associated with highly valued assets.

1. Introduction

When economists model the behavior of individual investors, they typically assume that these investors derive utility only from consumption or from total wealth. In this paper, we study the possibility that investors also derive utility from another source, namely from *realized gains and losses* on assets that they own. Suppose, for example, that an investor buys shares of a stock and then, a few months later, sells them. We consider a model in which he receives a burst of utility right then, at the moment of sale. The amount of utility depends on the size of the gain or loss realized—on the difference between the sale price and the purchase price—and is positive if the investor realizes a gain, and negative otherwise. This source of utility, which we label "realization utility," is not new to our paper: other authors also discuss it. Our contribution is to offer a comprehensive analysis of its implications for trading behavior and for asset prices.

Why might an investor derive utility from realizing a gain or loss? We think that realization utility is a consequence of two underlying cognitive processes. The first has to do with how people think about their investing history. Under this view, people do not think about their investing history purely in terms of the return they have earned on their portfolio. Rather, they often think about it as a series of investing episodes, each one defined by three things: the name of the investment, the purchase price, and the sale price. "I bought IBM at $80 and sold it at $100" might be one such episode. "We bought our house for $260,000 and sold it for $320,000" might be another.

The second cognitive process that, in our view, underlies realization utility has to do with how people *evaluate* their investing episodes. We suspect that many investors use a simple heuristic to guide their trading, one that says: "Selling a stock at a gain relative to purchase price is a good thing—it is what successful investors do." After all, an investor who buys a number of stocks in sequence and manages to realize a gain on all of them *does* end up with more money than he had at the start. The flip side of the same heuristic says: "Selling a stock at a loss is a bad thing—it is what unsuccessful investors do." Indeed, an investor who buys a number of stocks in sequence and realizes a loss on all of them *does* end up with less money than he had at the start.

In summary, an investor feels good when he sells a stock at a gain because, by selling, he is creating what he views as a *positive* investing episode. Conversely, he feels bad when he sells a stock at a loss because, by selling, he is creating what he views as a *negative* investing episode.

We do not expect realization utility to be important for all investors or in all circum-

2

stances. For example, we expect it to matter more for individual investors than for institutional investors who, as trained professionals, are more likely to think about their investing history in terms of overall portfolio return than as a series of investing episodes. Also, since realization utility depends on the difference between sale price and purchase price, it is likely to play a larger role when the purchase price is more salient. It may therefore be more relevant to the trading of individual stocks or to the sale of real estate than to the trading of mutual funds: the purchase price of a stock or of a house is typically more salient than that of a fund.

In our view, the idea that some investors derive utility directly from realizing gains and losses is a plausible one. But in order to claim that realization utility is a significant driver of investor behavior, we cannot appeal to mere plausibility. To make a more convincing case, we need to build a model of realization utility and then see if the model explains a range of facts and leads to new predictions that can be tested and confirmed.

In this paper, we take up this challenge. We construct a model of realization utility, discuss its predictions, and show that it can shed light on a number of empirical facts. We start with a partial equilibrium framework but also show how realization utility can be embedded in a full equilibrium model. This allows us to make predictions not only about trading behavior but also about prices.

Our partial equilibrium model is an infinite horizon model in which, at each moment, an investor allocates his wealth either to a risk-free asset or to one of a number of stocks. If the investor sells his holdings of a stock, he receives a burst of utility based on the size of the gain or loss realized and pays a proportional transaction cost. He also faces the possibility of a random liquidity shock: if such a shock occurs, he must immediately sell his asset holdings and exit the asset markets. At each moment, the investor makes his allocation decision by maximizing the discounted sum of expected future utility flows. In our baseline model, we assume a linear functional form for realization utility. Later, we also consider a piecewise-linear specification.

We find that, under the optimal strategy, an investor who is holding a position in a stock will voluntarily sell this position only if the stock price rises sufficiently far above the purchase price. We look at how this "liquidation point" at which the investor sells depends on the expected stock return, the standard deviation of the stock return, the time discount rate, the transaction cost, and the likelihood of a liquidity shock.

The model has a number of interesting implications. One of the more striking is that, even if realization utility has a linear or concave functional form, the investor can be risk

3

seeking: all else equal, his initial value function can be an increasing function of the standard deviation of stock returns. The intuition is straightforward. A highly volatile stock offers the chance of a large gain which the investor can enjoy realizing. Of course, it may also drop a lot in value; but in that case, the investor will simply postpone selling the stock until he is forced to sell by a liquidity shock. Any realized loss therefore lies in the distant, discounted future and does not scare the investor very much at the time of purchase. Overall, then, the investor may prefer more volatility to less.

We use our model to link realization utility to a number of financial phenomena. Among the applications we discuss are the disposition effect (Shefrin and Statman, 1985; Odean, 1998), the subpar trading performance of individual investors (Barber and Odean, 2000; Barber et al., 2009), the higher volume of trade in bull markets than in bear markets (Stein, 1995; Statman, Thorley, and Vorkink, 2006; Griffin, Nardari, and Stulz, 2007), the effect of historical highs on the propensity to sell (Grinblatt and Keloharju, 2001), the individual investor preference for volatile stocks (Kumar, 2009), the low average return of volatile stocks (Ang et al., 2006), and the heavy trading associated with highly valued assets—as, for example, in the case of U.S. technology stocks in the late 1990s (Hong and Stein, 2007).

Of these applications of realization utility, the most obvious is the disposition effect, the greater propensity of individual investors to sell stocks that have risen in value, rather than fallen in value, since purchase. In combination with a sufficiently positive time discount rate, realization utility generates a strong disposition effect: the investor in our model voluntarily sells a stock only if it is trading at a gain relative to purchase price.

While the link between realization utility and the disposition effect is clear, we emphasize that realization utility is not a "relabeling" of the disposition effect. On the contrary, it is just one of a number of possible theories of the disposition effect and can be distinguished from other theories through carefully constructed tests. For example, another theory of the disposition effect, one that has nothing to do with realization utility, is that investors have an irrational belief in mean-reversion. Later in the paper, we discuss an experiment that can distinguish this view from the realization utility view.

Our other applications are more subtle. For example, our model predicts that individual investors—the investor group that is more likely to think in terms of realization utility—will have a much greater propensity to sell a stock once its price moves above its historical high. Imagine a stock that rises to a high of \$45, falls, and then rises again, passing its previous high of \$45 and continuing upwards. Our model predicts that there will be relatively little selling as the stock approaches \$45 for the second time—any realization utility investors with

4

liquidation points of $45 or lower *will have sold already* when the stock first approached $45—but once the stock moves above the historical high of $45, realization utility investors with liquidation points higher than $45 will start to sell. In line with the evidence of Grinblatt and Keloharju (2001), then, our model predicts that historical highs will have a sharp effect on individual investors' propensity to sell.

The idea that people derive utility from gains and losses rather than from final wealth levels was first proposed by Markowitz (1952), but is particularly associated with Kahneman and Tversky (1979): it is a central element of their prospect theory model of decision-making. Finance researchers have typically taken Kahneman and Tversky's message to be that they should study models in which investors derive utility from *paper* gains and losses. Benartzi and Thaler (1995), for example, assume that investors derive utility from fluctuations in their financial wealth, while Barberis, Huang, and Santos (2001) and Barberis and Huang (2001) assume that they derive utility from fluctuations in the value of their stock market holdings or in the value of specific stocks that they own.

The idea that people might derive utility from *realized* gains and losses has received much less attention. The concept first appears in Shefrin and Statman (1985). Among several other contributions, these authors point out, with the help of a numerical example, that if an investor derives utility from realized gains and losses and has a utility function that, as in prospect theory, is concave over gains and convex over losses, then he will exhibit a disposition effect.

Shefrin and Statman (1985) justify their emphasis on realized gains and losses by reference to "mental accounting," a term used to describe how people think about, organize, and evaluate their financial transactions. In their view, when an investor sells a stock, he is closing a mental account that was opened when he first bought the stock. The moment of sale is therefore a natural time at which to evaluate the transaction: a realized gain is seen as a good outcome and a realized loss as a poor outcome. Realized gains and losses thereby become carriers of utility in their own right. Although described using different language, this motivation for realization utility is similar to our own.[1]

More recently, Barberis and Xiong (2009) use a two-period model to study the trading behavior of an investor who derives utility from realized gains and losses with a utility function that is concave over gains and convex over losses. They observe that, consistent with Shefrin and Statman (1985), the investor often exhibits a disposition effect. They

---

[1]Other authors also discuss realization utility. For example, Thaler (1999) writes that "one clear intuition is that a realized loss is more painful than a paper loss. When a stock is sold, the gain or loss has to be 'declared' both to the tax authorities and to the investor (and spouse)."

do not study any other implications of realization utility, nor do they link it to any other applications.[2]

In this paper, we offer a more comprehensive analysis of realization utility. We construct a richer model—an infinite horizon model that allows for transaction costs and a stochastic liquidity shock. We derive an analytical solution for the investor's optimal trading strategy. We show how realization utility can be incorporated into both a model of trading behavior and a model of asset pricing. We document several basic implications of realization utility. And we discuss many potential applications, rather than just one.

In Section 2, we present a partial equilibrium model of realization utility, one that also assumes a linear functional form for the realization utility term. In Section 3, we use a piecewise-linear functional form. In Section 4, we show how realization utility can be embedded in a model of asset prices. Section 5 discusses a range of applications and testable predictions, while Section 6 concludes.

2. A model of realization utility

Before presenting our model, we briefly note two of our assumptions. First, we assume that realization utility is defined at the level of an individual asset—a stock, a house, or a mutual fund, say. Realization utility is triggered by the act of selling. But when an investor makes a sale, he is selling a specific asset. It is therefore natural to define realization utility at the level of this asset. This assumption has little bite in our baseline model because, in this model, the investor holds at most one risky asset at any time. However, it becomes more important when we discuss an extension of our model in which the investor can hold several risky assets simultaneously.

A second assumption concerns the functional form for realization utility. In this section, we use a linear functional form so as to show that we do not need elaborate specifications in order to draw interesting implications out of realization utility. In Section 3, we also consider a piecewise-linear functional form.

We work in an infinite horizon, continuous time framework. An investor starts at time 0 with wealth W_0. At each time $t \geq 0$, he has the following investment options: a risk-free asset, which offers a constant continuously compounded return of r; and N risky assets indexed by $i \in \{1, \ldots, N\}$. The most natural application of our model is to understanding

---

[2]Barberis and Xiong (2009) do not say very much about realization utility because it is not their main focus. Their paper is primarily about the trading behavior of an investor who derives prospect theory utility from *paper* gains and losses.

how individual investors trade stocks in their brokerage accounts. We therefore often refer to the risky assets as stocks.

The price of stock i, $S_{i,t}$, follows

$$\frac{dS_{i,t}}{S_{i,t}} = (r + \mu)dt + \sigma dZ_{i,t}, \tag{1}$$

where $Z_{i,t}$ is a Brownian motion and where, for $i \neq j$, $dZ_{i,t}$ and $dZ_{j,t}$ may be correlated. In the interval between t and $t + dt$, stock i also pays a dividend flow of

$$D_{i,t}dt = \alpha S_{i,t}dt. \tag{2}$$

The stock's expected excess return—throughout the paper, "excess" means over and above the risk-free rate—is therefore $\alpha + \mu$: the dividend yield α plus the expected excess capital gain μ. For now, we assume that each of α, μ, and σ is the same for all stocks.

The dividends $D_{i,t}$ do not play a significant role in the partial equilibrium analysis in Sections 2 and 3. The only reason we introduce them is because, as we will see in Section 4, they make it easier to embed realization utility in a full equilibrium framework. To prevent the dividends from unnecessarily complicating the analysis, we make the following assumptions about them: that the investor consumes them; and that he receives linear consumption utility

$$v(c) = \beta c \tag{3}$$

from doing so, where β determines the importance of consumption utility relative to the second source of utility that we introduce below.

We assume that, at each time t, the investor either allocates all of his wealth to the risk-free asset or all of his wealth to one of the stocks; for simplicity, no other allocations are allowed. Therefore, over any interval of time during which the investor maintains a position in one particular asset, his wealth W_t evolves according to

$$\frac{dW_t}{W_t} = rdt + \sum_{i=1}^{N}(\mu dt + \sigma dZ_{i,t})\theta_{i,t}, \tag{4}$$

where $\theta_{i,t}$ takes the value one if he is holding stock i at time t, and zero otherwise. Note that, if $\theta_{i,t} = 1$ for some i and t, then $\theta_{j,t} = 0$ for all $j \neq i$. We also suppose that, if the investor sells his position in a stock at time t, he pays a proportional transaction cost kW_t, $0 \leq k < 1$.

An important variable in our model is B_t. This variable, which is formally defined only if the investor is holding a stock at time t, measures the cost basis of the stock position, in

7

other words, the reference point relative to which the investor computes his realized gain or loss. One possible definition of the cost basis is the amount of money the investor put into the time t stock position at the time he bought it. This is the definition we use, with one adjustment. We take the cost basis to be the amount of money the investor put into the stock position at the time he bought it, *scaled up* by the risk-free return between the time of purchase and time t, so that

$$B_t = W_s e^{r(t-s)}, \tag{5}$$

where $s \leq t$ is the moment at which the time t stock position was purchased. This definition is tractable and may be more realistic than the alternative that sets the cost basis equal to the original purchase price: the investor may only think of an investing episode as a positive one if the capital gain exceeds what he could have earned by investing in the risk-free asset.

The key feature of our model is that the investor derives utility from realizing a gain or loss. If, at time t, he moves his wealth from a stock into the risk-free asset or into another stock, he receives a burst of utility given by

$$u((1-k)W_t - B_t). \tag{6}$$

The argument of the utility term is the realized gain or loss: the investor's wealth at the moment of sale net of the transaction cost, $(1-k)W_t$, minus the cost basis of the stock investment B_t. Throughout this section, we use the linear functional form

$$u(x) = x. \tag{7}$$

We emphasize that the investor only receives the burst of utility in (6) if he moves his wealth from a stock into the risk-free asset or into *another* stock. If he sells a stock and then immediately puts the proceeds back into the *same* stock, he derives no realization utility from the sale, nor is the cost basis affected. Realization utility is associated with the end of an investing episode. It is hard to argue that the sale of a stock represents the end of an episode if, after selling the stock, the investor immediately buys it back.

We assume that the investor does not incur a transaction cost if he sells the risk-free asset. If we measure the cost basis for this asset in the same way as for a stock, it follows that the realized gain or loss from selling the risk-free asset is always zero. The investor therefore receives realization utility only when he sells a stock, not when he sells the risk-free asset.

The investor also faces the possibility of a random liquidity shock whose arrival is governed by a Poisson process with parameter ρ. If a shock occurs, the investor immediately

sells his holdings, exits the asset markets, and, if he was holding a stock at the time of the shock, receives the burst of utility in (6). We think of this shock as capturing a sudden consumption need that forces the investor to draw on the funds in his brokerage account. We include it because it ensures, as is reasonable, that the investor cares not only about realized gains and losses but also about *paper* gains and losses. It also gives us a way of varying the investor's horizon: when ρ is high, the investor effectively has a short horizon; when it is low, he has a long horizon.

At each moment, the investor makes his allocation decision by maximizing the discounted sum of expected future utility flows. Suppose that, at time t, his wealth is allocated to a stock. His value function then depends on two things: on the current value of his position, W_t, and on the cost basis of the position, B_t. We therefore denote it as $V(W_t, B_t)$. Since the utility functions in (3) and (7) are homogeneous of degree one, and since the prices of the risky assets all follow geometric Brownian motions, the value function must also be homogeneous of degree one, so that, for $\zeta > 0$,

$$V(\zeta W_t, \zeta B_t) = \zeta V(W_t, B_t). \tag{8}$$

Now suppose that, for some positive W,

$$V(W, W) \geq 0. \tag{9}$$

Note that $V(W, W)$ is the value function that corresponds to investing wealth W in a stock *now*, so that current wealth and the cost basis are both equal to W. Since $V(W_t, B_t)$ is homogeneous of degree one, if (9) holds for some positive W, then it holds for all positive W. Later, we will compute the range of parameter values for which (9) holds. For now, we note that, so long as the time discount rate δ exceeds the risk-free return r, condition (9) implies two things. First, it implies that, at time 0, the investor allocates his wealth to one of the N stocks: since the risk-free asset generates no utility flows, he allocates to a stock as early as possible. Second, and using the same logic, condition (9) implies that, if, at any time $t > 0$, the investor sells his holdings of a stock, he will then immediately use the proceeds to buy another stock.

We can now formulate the investor's decision problem. Suppose that, at time t, the investor is holding stock i. Let τ' be the random future time at which a liquidity shock occurs. Then, at time t, the investor solves

$$
\begin{aligned}
V(W_t, B_t) \quad = \quad \max_{\tau \geq t} \quad E_t \Big\{ & \int_t^{\min\{\tau, \tau'\}} e^{-\delta(s-t)} v(D_{i,s}) ds \\
& + e^{-\delta(\tau-t)} [u((1-k)W_\tau - B_\tau) + V((1-k)W_\tau, (1-k)W_\tau)] I_{\{\tau < \tau'\}} \\
& + e^{-\delta(\tau'-t)} u((1-k)W_{\tau'} - B_{\tau'}) I_{\{\tau \geq \tau'\}} \Big\},
\end{aligned} \tag{10}
$$

9

subject to (3), (4), (5), and (7). $I_{\{\}}$ is an indicator function that takes the value one if the condition in the curly brackets is met, and zero otherwise. To ensure that the investor does not hold his time 0 stock position forever, without selling it, we impose the following parameter restriction, which, in words, requires that the expected excess capital gain is not too high:

$$\mu < (\rho + \delta - r)\left(1 - \frac{k}{\delta - r}(\rho + \frac{\alpha\beta}{1-k})\right). \tag{11}$$

Note that this implies $\mu < \rho + \delta - r$, a simpler condition that we will sometimes also use.

To understand the formulation in (10), note that the investor's problem is to choose the optimal time τ, a random time in the future, at which to realize the gain or loss in his stock holdings. Suppose first that $\tau < \tau'$, so that the investor voluntarily sells the stock before a liquidity shock arrives. In this case, the investor receives a burst of utility $u((1-k)W_\tau - B_\tau)$ when he sells at time τ; and a cash balance of $(1-k)W_\tau$ which he immediately invests in another stock. If $\tau \geq \tau'$, however, the investor is forced out of the stock market by a liquidity shock and receives realization utility $u((1-k)W_{\tau'} - B_{\tau'})$ from the gain or loss at the moment of exit. Finally, while holding the stock, the investor receives a continuous stream of dividends.

The proposition below presents the solution to the decision problem in (10). It states that if the investor buys a stock, his optimal strategy is to sell it voluntarily only if its price rises a sufficient amount above the purchase price. The variable

$$g_t = \frac{W_t}{B_t} \tag{12}$$

—in words, the value of the stock position the investor is holding at time t relative to its cost basis—plays an important role in the solution. To simplify the statement of the proposition, we define

$$\delta' \equiv \delta - r. \tag{13}$$

As we will see, the investor's behavior does not depend on δ and r separately, but only on the difference between them. We sometimes refer to δ' as the "effective" discount rate and assume throughout that $\delta' > 0$. The proof of the proposition is in the Appendix.

Proposition 1. *Unless forced to exit the stock market by a liquidity shock, an investor with the decision problem in (10) will sell his holdings of a stock if the gain $g_t = W_t/B_t$ reaches a liquidation point $g_t = g_* \geq 1$. If the transaction cost k is positive, then $g_* > 1$. The value*

10

function is $V(W_t, B_t) = B_t U(g_t)$, *where*[3]

$$U(g_t) = \begin{cases} a g_t^{\gamma_1} + \frac{\alpha\beta + \rho(1-k)}{\rho+\delta'-\mu} g_t - \frac{\rho}{\rho+\delta'} & \text{if} \quad g_t \in (0, g_*) \\ \\ (1-k)(1+U(1))g_t - 1 & \text{if} \quad g_t \in [g_*, \infty) \end{cases}, \qquad (14)$$

where

$$\gamma_1 = \frac{1}{\sigma^2} \left[\sqrt{\left(\mu - \frac{1}{2}\sigma^2\right)^2 + 2(\rho + \delta')\sigma^2} - \left(\mu - \frac{1}{2}\sigma^2\right) \right] > 0 \qquad (15)$$

and

$$a = \frac{\delta'}{g_*^{\gamma_1}(\gamma_1 - 1)(\rho + \delta')}. \qquad (16)$$

The liquidation point g_* *is the unique root, in the range* $[1, \infty)$, *of*

$$(\gamma_1 - 1)\left(1 - \frac{k(\rho+\delta')(\rho + \frac{\alpha\beta}{1-k})}{\delta'(\rho + \delta' - \mu)}\right) g_*^{\gamma_1} - \frac{\gamma_1}{1-k} g_*^{\gamma_1 - 1} + 1 = 0. \qquad (17)$$

In summary, the optimal strategy takes one of two forms. If the model parameters are such that $U(1) \geq 0$, where $U(1)$ is the value function per unit wealth from buying a stock at time 0—equivalently, if condition (9) holds—the investor buys a stock at time 0 and voluntarily sells it only if it reaches a sufficiently high liquidation point, at which time he immediately invests the proceeds in another stock, and so on. In particular, the investor never voluntarily sells a stock at a loss. If, on the other hand, $U(1) < 0$, the investor allocates his wealth to the risk-free asset at time 0 and keeps it there until a liquidity shock arrives.[4]

For expositional simplicity, we have assumed that the investor holds at most one stock at any time. However, Proposition 1 can also tell us how the investor trades in a setting where he holds several stocks simultaneously. Suppose that, at time 0, he spreads his wealth across a number of stocks. Suppose also, as is natural in the case of realization utility, that he derives utility *separately* from the realized gain or loss on each stock. Finally, suppose that if a liquidity shock occurs, the investor sells all of his holdings and exits the asset markets. Under these assumptions, the investor's decision problem is "separable" across the different stocks he is holding and the solution to (10) in Proposition 1 describes how he trades each one of his stocks.

---

[3]Since $g_* \geq 1$, the term $U(1)$ which appears in the second row of Eq. (14) can be obtained from the first row of the equation. It equals $a + (\alpha\beta + \rho(1-k))/(\rho + \delta' - \mu) - \rho/(\rho + \delta')$.

[4]To be clear, if $g_* = 1.05$, say, the investor sells his holdings of a stock once the value of the position is 5% higher than the cost basis. Given the definition of the cost basis in (5), this means that the value of the position at the time of sale is *more* than 5% higher than it was at the time of purchase.

A corollary to Proposition 1—one that also holds for the piecewise-linear specification we consider in Section 3—is that, in this multiple-concurrent-stock extension of our basic model, the investor is indifferent to diversification. For example, he is indifferent between investing W_0 in just one stock at time 0 as compared to investing $W_0/2$ in each of two stocks at time 0. The time 0 value function for the first strategy, $W_0 U(1)$, is the same as the time 0 value function for the second strategy, namely $W_0 U(1)/2 + W_0 U(1)/2$.

2.1. Results

In this section, and again in Section 3, we draw out the implications of realization utility through two kinds of analysis. First, we compute the range of parameter values for which condition (9) holds, so that the investor is willing to buy a stock at time 0. Second, we look at how the liquidation point g_* and initial utility per unit wealth $U(1)$ depend on each of the model parameters. The first analysis therefore concerns the investor's buying behavior, and the second, his selling behavior. When assigning parameter values, we have in mind our model's most natural application, namely, the trading of stocks by individual investors.

The shaded area in the top graph in Fig. 1 shows the range of values of the expected excess stock return $\alpha + \mu$ and standard deviation of stock returns σ that satisfy $U(1) \geq 0$—in other words, condition (9)—so that the investor is willing to buy a stock at time 0, but also the restriction in (11), so that he sells the stock at a finite liquidation point.[5]

To create the graph, we assign values to δ', k, ρ, α, and β, and then search for values of μ and σ such that both $U(1) \geq 0$ and condition (11) hold. We set the transaction cost to $k = 0.005$ and the liquidity shock intensity ρ to 0.1, so that the probability of a shock over the course of a year is $1 - e^{-0.1} \approx 0.1$. We also set the dividend yield α to 0.015 and the consumption utility weight β to 1. Finally, we choose an effective discount rate of $\delta' = 0.08$ because, as we will see later, this generates a trading frequency similar to that observed in actual brokerage accounts.

The graph illustrates an interesting implication of realization utility, namely that the investor is willing to buy a stock with a *negative* expected excess return, so long as its standard deviation σ is sufficiently high. The intuition is straightforward. So long as σ is sufficiently high, even a negative expected excess return stock has a non-negligible chance of reaching the liquidation point g_*, at which time the investor can enjoy realizing a gain. Of course, more likely than not, the stock will perform poorly. However, since the investor does

---

[5]The unshaded area in the bottom-left of the graph corresponds to parameter values for which $U(1) < 0$, so that the investor does not buy a stock at time 0. The unshaded area in the right of the graph corresponds to parameter values that violate restriction (11).

not voluntarily realize losses, this will only bring him disutility in the event of a liquidity shock. Any realized loss therefore lies in the distant, discounted future and does not scare the investor very much at the time of purchase. Overall, then, investing in a stock with a low expected return can sometimes be better than investing in the risk-free asset.

Figs. 2 and 3 show how the liquidation point g_* and initial utility per unit wealth $U(1)$ depend on the parameters μ, σ, δ', k, and ρ. The graphs on the left side of each figure correspond to the liquidation point, and those on the right side, to initial utility. For now, we focus on the solid lines; we discuss the dashed lines in Section 3.

To construct the graphs, we start with a set of benchmark parameter values. We use the same benchmark values throughout the paper. Consider first the asset-level parameters α, μ, σ, and k. We assume a dividend yield α of 0.015, an expected excess capital gain on stocks of $\mu = 0.015$—note that this implies an expected excess stock return of $\alpha + \mu = 0.03$—a standard deviation of stock returns of $\sigma = 0.5$, and a transaction cost of $k = 0.005$. As for the investor-level parameters δ', ρ, and β, we use an effective time discount rate of $\delta' = 0.08$, a liquidity shock intensity of $\rho = 0.1$, and a consumption utility weight of $\beta = 1$. The graphs in Figs. 2 and 3 vary each of μ, σ, δ', k, and ρ in turn, keeping the other parameters fixed at their benchmark values.

The top-right graph in Fig. 2 shows that, as is natural, initial utility is increasing in the expected excess capital gain μ. The top-left graph shows that the liquidation point is also increasing in μ: if a stock that is trading at a gain has a high expected return, the investor is tempted to hold on to it rather than to sell it and incur a transaction cost.

The middle-right graph illustrates an important implication of realization utility: that, as stock return volatility goes up, initial utility also goes up. Put differently, even though realization utility has a linear functional form, the investor is risk *seeking*. The intuition for this parallels the intuition for why the investor is sometimes willing to buy a stock with a low expected return. The more volatile a stock is, the more likely it is to reach its liquidation point, at which time the investor can enjoy realizing a gain. Of course, a volatile stock may also decline a lot in value. But the investor does not voluntarily realize losses and so will only experience disutility in the event of a liquidity shock. Any realized loss therefore lies in the distant, discounted future and does not scare the investor very much at the time of purchase. Overall, then, the investor prefers more volatility to less.[6] A similar intuition explains why, in the middle-left graph, the liquidation point is increasing in volatility.

---

[6]In mathematical terms, this prediction is related to the fact that, while instantaneous utility is linear, the value function $U(g_t)$ in (14) is convex: since, from (11), $\mu < \rho + \delta'$, we have $\gamma_1 > 1$ and $a > 0$, which, in turn, imply the convexity of $U(\cdot)$.

The trading patterns we have just described—the buying of low expected return stocks and the preference for volatile stocks—are not behaviors that we associate with sophisticated investors. We emphasize, however, that our model is not a model of sophisticated investors. It is a model of unsophisticated investors—specifically, of investors who use a simple heuristic to guide their trading, one that says that selling an asset at a gain is a good thing and that selling an asset at a loss is a bad thing. What Figs. 1 and 2 demonstrate is that an investor who thinks in these terms can be drawn into stocks with low expected returns and high volatility. We discuss some evidence consistent with this prediction in Section 5.[7]

The bottom-left graph in Fig. 2 shows that when the investor discounts the future more heavily, the liquidation point falls. An investor with a high discount rate is impatient and therefore wants to realize gains sooner rather than later.

The top graphs in Fig. 3 show how the liquidation point and initial utility depend on the transaction cost k. As expected, a higher transaction cost lowers time 0 utility. It also increases the liquidation point: if it is costly to sell a stock, the investor waits longer before doing so.

What happens when there is no transaction cost? The top-left graph in Fig. 3 suggests that, in this case, the liquidation point is $g_* = 1$. It is straightforward to check that when $k = 0$, (17) is indeed satisfied by $g_* = 1$, so that the investor realizes all gains immediately. In other words, in our model, it is the transaction cost that stops the investor from realizing all gains as soon as they appear.

The bottom graphs in Fig. 3 show how the liquidation point and initial utility depend on ρ, the intensity of the liquidity shock. The liquidation point depends on ρ in a non-monotonic way. There are two forces at work here. As the liquidity shock intensity ρ goes up, the liquidation point initially falls. One reason the investor delays realizing a gain is the transaction cost that a sale entails. For $\rho > 0$, however, the investor knows that he will be forced out of the stock market at some point. The present value of the transaction costs he expects to pay is therefore lower than in the absence of liquidity shocks. As a result, he is willing to realize gains sooner.

At higher levels of ρ, however, another factor makes the investor more patient. If he is holding a stock with a gain, he is reluctant to exit the position because he will then have to invest the proceeds in another stock, which might do poorly and which he might be forced

---

[7]For the case of linear realization utility, the predictions that the investor will be willing to buy stocks with low expected returns and that he will be risk seeking are robust to changes in the model parameters. In the next section, however, we will see that when the investor is more sensitive to realized losses than to realized gains, these predictions do not always hold.

to sell at a loss by a liquidity shock. This factor pushes the liquidation point back up.

The bottom-right graph shows that as the liquidity shock intensity rises, initial utility falls. A high intensity ρ makes it more likely that in the near future, the investor will be forced to exit the stock market with a painful loss.

Several of the implications of realization utility that we have described can also be obtained in a two-period version of our model. However, our infinite horizon framework has at least one advantage. In an infinite horizon model, the structure of the optimal trading strategy is simpler than in a two-period model: the investor either holds the risk-free asset or else buys a series of stocks in sequence, selling each one whenever it reaches a fixed liquidation point. The reason for this simple structure is that in the infinite horizon model, the environment is stationary: the value function does not depend explicitly on time, t. In a two-period model, the environment is non-stationary and so the optimal trading strategy, while similar to that in our model, has a more complex structure.

We have also studied an extension of our model in which the value of the dividend yield α, the expected excess capital gain μ, and the standard deviation of returns σ differ across stocks. In this case, the investor follows a strategy that is similar to the one described above, but that is restricted to a subset of the available stocks. Specifically, for each stock i, the investor computes $V_i(W, W)$, the value function from investing wealth W in stock i today. Suppose that stock j, with parameter values α_j, μ_j, and σ_j, maximizes $V_i(W, W)$ across all stocks; and suppose also that there are several stocks, which together comprise a set M, that have the same parameter values as stock j. Then, so long as $V_j(W, W) \geq 0$, the investor allocates his wealth to a stock drawn from M at time 0, sells it when it reaches the liquidation point specified in Proposition 1, and then immediately reinvests the proceeds in another stock drawn from M, and so on.

Fig. 2 tells us something about the characteristics of the stocks in the agent's preferred set M: a stock is more likely to be in M, the higher its expected excess capital gain μ and the higher its standard deviation σ. Realization utility therefore has implications not only for an investor's selling behavior, but also for his buying behavior.

3. The case of piecewise-linear utility

In Section 2, we took the functional form for realization utility $u(\cdot)$ to be linear. However, in reality, investors may be more sensitive to realized losses than to realized gains. We

therefore now look at what happens when $u(\cdot)$ is piecewise-linear rather than linear:

$$u(x) = \begin{cases} x & \text{if} \quad x \geq 0 \\ \lambda x & \text{if} \quad x < 0 \end{cases}, \quad \lambda > 1, \tag{18}$$

where λ determines the relative sensitivity to realized losses as opposed to realized gains.[8]

The investor's decision problem is now

$$\begin{aligned} V(W_t, B_t) &= \max_{\tau \geq t} E_t\{\int_t^{\min\{\tau,\tau'\}} e^{-\delta(s-t)} v(D_{i,s}) ds \\ &\quad + e^{-\delta(\tau-t)}[u((1-k)W_\tau - B_\tau) + V((1-k)W_\tau, (1-k)W_\tau)]I_{\{\tau<\tau'\}} \\ &\quad + e^{-\delta(\tau'-t)}u((1-k)W_{\tau'} - B_{\tau'})I_{\{\tau\geq\tau'\}}\}, \end{aligned} \tag{19}$$

subject to (3), (4), (5), and (18). This is the same as decision problem (10) in Section 2 except that $u(\cdot)$ is no longer linear but instead takes the form in (18).

In the Appendix, we prove:

Proposition 2. *Unless forced to exit the stock market by a liquidity shock, an investor with the decision problem in (19) will sell his holdings of a stock if the gain $g_t = W_t/B_t$ reaches a liquidation point $g_t = g_* \geq 1$. If the transaction cost k is positive, then $g_* > 1$. The value function is $V(W_t, B_t) = B_t U(g_t)$, where*

$$U(g_t) = \begin{cases} bg_t^{\gamma_1} + \frac{\alpha\beta+\rho\lambda(1-k)}{\rho+\delta'-\mu}g_t - \frac{\rho\lambda}{\rho+\delta'} & \text{if} \quad g_t \in \left(0, \frac{1}{1-k}\right) \\ c_1 g_t^{\gamma_1} + c_2 g_t^{\gamma_2} + \frac{\alpha\beta+\rho(1-k)}{\rho+\delta'-\mu}g_t - \frac{\rho}{\rho+\delta'} & \text{if} \quad g_t \in \left[\frac{1}{1-k}, g_*\right) , \\ (1-k)g_t(1+U(1)) - 1 & \text{if} \quad g_t \in [g_*, \infty) \end{cases} \tag{20}$$

where γ_1 is defined in (15), where

$$\gamma_2 = -\frac{1}{\sigma^2}\left[\sqrt{\left(\mu - \frac{1}{2}\sigma^2\right)^2 + 2(\rho+\delta')\sigma^2} + \left(\mu - \frac{1}{2}\sigma^2\right)\right] < 0, \tag{21}$$

---

[8]It is not clear whether a piecewise-linear form is more reasonable than a linear one. There is, of course, the well-known concept of "loss aversion," but this is the idea that people are more sensitive to *wealth* losses than to wealth gains, in other words, more sensitive to *paper* losses than to paper gains. It is the premise of this paper that utility from realized gains and losses is distinct from utility from paper gains and losses and that it may have different psychological roots. Even if people are more sensitive to paper losses than to paper gains, it does not necessarily follow that they are also more sensitive to realized losses than to realized gains.

and where b, c_1, c_2, and g_* are determined from

$$c_2 = \frac{(\lambda - 1)\,\rho(1 - k)^{\gamma_2}\,(\mu\gamma_1 - \rho - \delta')}{(\gamma_1 - \gamma_2)\,(\rho + \delta' - \mu)\,(\rho + \delta')} \tag{22}$$

$$(\gamma_1 - 1)\,c_1 g_*^{\gamma_1} + (\gamma_2 - 1)\,c_2 g_*^{\gamma_2} = \frac{\delta'}{\rho + \delta'} \tag{23}$$

$$c_1 \left(\frac{1}{1-k}\right)^{\gamma_1} + c_2 \left(\frac{1}{1-k}\right)^{\gamma_2} = b\left(\frac{1}{1-k}\right)^{\gamma_1} + \frac{(\lambda - 1)\mu\rho}{(\rho + \delta' - \mu)(\rho + \delta')} \tag{24}$$

$$c_1 g_*^{\gamma_1} + c_2 g_*^{\gamma_2} + \frac{k\alpha\beta + (1 - k)\,(\mu - \delta')}{\rho + \delta' - \mu} g_* + \frac{\delta'}{\rho + \delta'} = (1 - k)g_* \left(b + \frac{\rho\lambda(\mu - k\rho - k\delta')}{(\rho + \delta')(\rho + \delta' - \mu)}\right) \tag{25}$$

Specifically, given values for the asset-level parameters α, μ, σ, and k, and for the investor-level parameters δ', ρ, λ, and β, we first use (22) to find c_2; we then obtain c_1 from (23); we then use (24) to find b; finally, (25) allows us to solve for the liquidation point g_*.

3.1. Results

The shaded area in the lower graph in Fig. 1 shows the range of values of the expected excess stock return $\alpha + \mu$ and standard deviation of stock returns σ for which the investor is willing to buy a stock at time 0—in other words, condition (9) is satisfied—but also to sell the stock at a finite liquidation point. We set the asset-level parameters α and k to their benchmark values from before, namely 0.015 and 0.005, respectively; and we set the investor-level parameters δ', ρ, and β to their benchmark values of 0.08, 0.1, and 1, respectively. Finally, we assign λ the benchmark value of 1.5.

Relative to the upper graph—the graph for the Section 2 model with *linear* realization utility—we see that the investor is now more reluctant to invest in a stock with a negative expected excess return. For a realization utility investor, the problem with investing in such a stock is that it raises the chance that he will be forced, by a liquidity shock, to make a painful exit from a losing position. A high sensitivity to losses makes this prospect all the more unappealing. The investor therefore only invests in a negative expected excess return stock if it is highly volatile, so that it at least offers a non-negligible chance of a sizeable gain that he can enjoy realizing.

When $\lambda > 1$, the prediction that the investor will be willing to invest in a stock with a negative expected excess return depends heavily on the parameters ρ, λ, and δ'. If the liquidity shock intensity or the sensitivity to losses rise significantly above their benchmark values, or if the discount rate falls significantly below its benchmark value, the investor will no longer be willing to buy a negative expected excess return stock, whatever its volatility.

The graphs in Fig. 4 show how the liquidation point g_* and initial utility per unit wealth

$U(1)$ depend on the sensitivity to losses λ. These graphs vary λ while maintaining

$$
\begin{aligned}
(\alpha, \mu, \sigma, k) &= (0.015, 0.015, 0.5, 0.005) \\
(\delta', \rho, \beta) &= (0.08, 0.1, 1).
\end{aligned} \tag{26}
$$

In the left graph, we see that the more sensitive the investor is to losses, the higher the liquidation point: a higher λ means that the investor is more reluctant to sell a stock at a gain, because if he does, he will have to invest the proceeds in a new stock, which might go down and which he might be forced to sell at a loss by a liquidity shock. The right graph shows that, as the sensitivity to losses goes up, initial utility falls: a high λ means that the investor may be forced, by a liquidity shock, to make an especially painful exit from a losing position.

The dashed lines in Fig. 2 show how the liquidation point g_* and initial utility $U(1)$ depend on μ, σ, and δ' when the investor is more sensitive to losses than to gains. Here, we vary each of μ, σ, and δ' in turn, keeping the other parameters fixed at their benchmark values

$$
\begin{aligned}
(\alpha, \mu, \sigma, k) &= (0.015, 0.015, 0.5, 0.005) \\
(\delta', \rho, \lambda, \beta) &= (0.08, 0.1, 1.5, 1).
\end{aligned} \tag{27}
$$

By comparing the dashed lines to the solid lines—the lines that correspond to linear realization utility—we see that, for our benchmark parameter values, allowing for greater sensitivity to losses preserves the qualitative relationship between g_* and $U(1)$ on the one hand, and μ, σ, and δ' on the other.

The dashed line in the middle-right graph of Fig. 2 deserves particular attention. It shows that, for the benchmark values in (27), initial utility $U(1)$ is still increasing in stock return volatility σ. Put differently, even though the functional form for realization utility is now concave, the investor is still risk seeking. However, when $\lambda > 1$, this prediction is sensitive to the values of ρ, λ, and δ'. If the sensitivity to losses or the liquidity shock intensity rise significantly, or if the discount rate falls significantly, the prediction is reversed: initial utility becomes a decreasing function of σ and the investor is risk averse, not risk seeking.

It is worth emphasizing the crucial role that the discount rate δ' plays in determining whether the investor is risk seeking or risk averse, and whether he is willing to buy stocks with low expected returns. Roughly speaking, buying a stock offers the investor either a short-term realized gain, should the stock perform well, or a long-term realized loss, should the stock perform poorly. The more impatient the investor is, the more he focuses on the

18

short-term gain as opposed to the long-term loss. As a result, he is more likely to be risk seeking and to invest in stocks with low expected returns.[9]

4. An asset pricing model

In Sections 2 and 3, we studied realization utility in a partial equilibrium model of trading behavior. In this section, we show how it can be embedded in an asset pricing model. We do not necessarily expect realization utility to have an impact on the prices of all stocks; it may, at most, affect the prices of stocks held and traded primarily by individual investors. Of course, the only way to know for sure is to derive the pricing implications of realization utility and to compare these predictions to the available facts.

Embedding non-standard preferences in a full equilibrium can be challenging. To make headway, we study the simplest possible model, one with *homogeneous* realization utility investors. Consider an economy with a risk-free asset and N risky stocks indexed by $i \in \{1, \ldots, N\}$. The risk-free asset is in perfectly elastic supply and earns a continuously compounded return of r. The risky stocks are in limited supply. The dividend process for stock i is

$$\frac{dD_{i,t}}{D_{i,t}} = (r + \mu_i)dt + \sigma_i dZ_{i,t}, \tag{28}$$

where $Z_{i,t}$ is a Brownian motion and where, for $i \neq j$, $dZ_{i,t}$ and $dZ_{j,t}$ may be correlated. The parameters μ_i and σ_i are constant over time but can vary across stocks.

The price of stock i at time t, $S_{i,t}$, is set in equilibrium. We hypothesize that

$$S_{i,t} = \frac{1}{\alpha_i} D_{i,t}, \tag{29}$$

where α_i will be determined later. By investing in stock i, an investor therefore receives the dividend stream $D_{i,t}$, which he consumes, and also the price fluctuation given by

$$\frac{dS_{i,t}}{S_{i,t}} = (r + \mu_i)dt + \sigma_i dZ_{i,t}. \tag{30}$$

The expected excess return of stock i is therefore $\alpha_i + \mu_i$.

The economy contains a continuum of realization utility investors. At each time $t \geq 0$, each investor must either allocate all of his wealth to the risk-free asset or all of his wealth

---

[9]We have also studied another extension of the model in Section 2, one that assumes hyperbolic, rather than exponential, discounting. We find that hyperbolic discounting has a significant effect on the trading behavior of an investor who is guided by realization utility. The more present-biased the investor is, the lower the liquidation point: a present-biased investor is impatient to realize gains. More generally, hyperbolic discounting is one way of thinking about the high discount rate δ required by condition (11).

to one of the stocks. We allow for transaction costs, liquidity shocks, and piecewise-linear utility. As noted above, the investors are homogeneous, so that δ', ρ, λ, and β are the same for all of them. Transaction costs, however, can differ across stocks. The transaction cost for stock i is k_i.

In this economy, the equilibrium conditions are

$$V_i(W, W) = 0, \qquad i = 1, \ldots, N, \tag{31}$$

where $V_i(W_t, B_t)$ is the value function for an investor whose wealth W_t is allocated to stock i and whose cost basis is B_t. In words, these conditions mean that an investor who is buying a stock is indifferent between allocating his wealth to that stock or to the risk-free asset.

Why are Eqs. (31) the appropriate equilibrium conditions? Note that, under the conditions in (31), we can clear markets at time 0 by assigning some investors to each stock and the rest to the risk-free asset. If, at any point in the future, some investors sell their holdings of stock i because of a liquidity shock, they immediately withdraw from the asset markets. If some investors sell their holdings of stock i because, for these investors, the stock has reached its liquidation point, the conditions in (31) mean that they are happy to then be assigned to the risk-free asset. Finally, the conditions in (31) mean that, if some investors do sell their holdings of stock i, whether because of a liquidity shock or because the stock reaches its liquidation point, we can reassign other investors from the risk-free asset to stock i, thereby again clearing the market in this stock.[10]

Formally, the decision problem for an investor holding stock i at time t is

$$
\begin{aligned}
V_i(W_t, B_t) \;=\; &\max_{\tau \geq t} \; E_t \Big\{ \int_t^{\min\{\tau, \tau'\}} e^{-\delta(s-t)} v(D_{i,s}) ds \\
&+ e^{-\delta(\tau-t)} u((1-k_i)W_\tau - B_\tau) I_{\{\tau < \tau'\}} + e^{-\delta(\tau'-t)} u((1-k_i)W_{\tau'} - B_{\tau'}) I_{\{\tau \geq \tau'\}} \Big\},
\end{aligned}
\tag{32}
$$

subject to (3), (5), (18), and

$$\frac{dW_s}{W_s} = (r + \mu_i)ds + \sigma_i dZ_{i,s}, \qquad t \leq s < \min\{\tau, \tau'\}, \tag{33}$$

where τ' is the random future time at which a liquidity shock arrives. This differs from the decision problem in (19) in that it imposes the market clearing condition (31): after selling his stock holdings at time τ, the investor's future value function is zero. We summarize the

---

[10]We assume here that whenever we need to reassign investors from the risk-free asset to one of the stocks, there are always enough investors holding the risk-free asset to make this possible. This can happen if, for example, investors who leave the asset markets because of a liquidity shock later re-enter.

solution to the decision problem in (32) in the following proposition. The proof is in the Appendix.

Proposition 3. *Unless forced to exit the stock market by a liquidity shock, an investor with the decision problem in (32) will sell his holdings of a stock if the gain $g_t = W_t/B_t$ reaches a liquidation point $g_t = g_* \geq 1$. If the transaction cost k_i is positive, then $g_* > 1$. The value function when holding stock i at time t is $V_i(W_t, B_t) = B_t U_i(g_t)$, where*

$$
U_i(g_t) = \begin{cases}
bg_t^{\gamma_1} + \frac{\alpha_i\beta + \rho\lambda(1-k_i)}{\rho+\delta'-\mu_i}g_t - \frac{\rho\lambda}{\rho+\delta'} & \text{if } g_t \in (0, \frac{1}{1-k_i}) \\[2ex]
c_1 g_t^{\gamma_1} + c_2 g_t^{\gamma_2} + \frac{\alpha_i\beta + \rho(1-k_i)}{\rho+\delta'-\mu_i}g_t - \frac{\rho}{\rho+\delta'} & \text{if } g_t \in [\frac{1}{1-k_i}, g_*), \\[2ex]
(1-k_i)g_t - 1 & \text{if } g_t \in [g_*, \infty)
\end{cases} \tag{34}
$$

where γ_1 and γ_2 are given by

$$
\gamma_1 = \frac{1}{\sigma_i^2}\left[\sqrt{\left(\mu_i - \frac{1}{2}\sigma_i^2\right)^2 + 2(\rho+\delta')\sigma_i^2} - \left(\mu_i - \frac{1}{2}\sigma_i^2\right)\right] > 0 \tag{35}
$$

$$
\gamma_2 = -\frac{1}{\sigma_i^2}\left[\sqrt{\left(\mu_i - \frac{1}{2}\sigma_i^2\right)^2 + 2(\rho+\delta')\sigma_i^2} + \left(\mu_i - \frac{1}{2}\sigma_i^2\right)\right] < 0, \tag{36}
$$

and where b, c_1, c_2, and g_* are determined from

$$
c_2 = \frac{(\lambda-1)\rho(1-k_i)^{\gamma_2}(\mu_i\gamma_1 - \rho - \delta')}{(\gamma_1-\gamma_2)(\rho+\delta'-\mu_i)(\rho+\delta')} \tag{37}
$$

$$
(\gamma_1-1)c_1 g_*^{\gamma_1} + (\gamma_2-1)c_2 g_*^{\gamma_2} = \frac{\delta'}{\rho+\delta'} \tag{38}
$$

$$
c_1\left(\frac{1}{1-k_i}\right)^{\gamma_1} + c_2\left(\frac{1}{1-k_i}\right)^{\gamma_2} = b\left(\frac{1}{1-k_i}\right)^{\gamma_1} + \frac{(\lambda-1)\mu_i\rho}{(\rho+\delta'-\mu_i)(\rho+\delta')} \tag{39}
$$

$$
c_1 g_*^{\gamma_1} + c_2 g_*^{\gamma_2} + \left(\frac{(1-k_i)(\mu_i-\delta'-\rho\lambda)}{\rho+\delta'-\mu_i} + \frac{\rho\lambda}{\rho+\delta'} - b\right)g_* = -\frac{\delta'}{\rho+\delta'}. \tag{40}
$$

The equilibrium expected excess return of stock i is $\alpha_i + \mu_i$. The parameter μ_i is the expected excess dividend growth rate and is exogenously given. To determine α_i, we require that the value function satisfies the condition in (31), namely $V_i(W, W) = 0$, or equivalently, $U_i(1) = 0$. The parameter α_i is therefore given by

$$
b + \frac{\alpha_i\beta + \rho\lambda(1-k_i)}{\rho+\delta'-\mu_i} - \frac{\rho\lambda}{\rho+\delta'} = 0. \tag{41}
$$

Since the parameters δ', ρ, λ, and β are constant across investors, α_i is constant over time, as assumed earlier.[11]

---

[11]In our model, it is the buyers of the risky assets, not the sellers, who set prices. In other words, the

In Section 5.2, we use the model described in this section to illustrate the effect of realization utility on asset prices. We emphasize that conditions (31) only describe an equilibrium when all investors in the economy have the same realization utility preferences. They do not describe an equilibrium when investors have heterogeneous realization utility preferences, nor when some investors have expected utility preferences defined only over consumption. We conjecture that in an economy with both expected utility and realization utility investors, the expected utility investors will partially—but only partially – attenuate any pricing effects caused by realization utility investors. The predictions of the model in this section should therefore hold more strongly among stocks traded by investors whose thinking is especially influenced by realization utility.

5. Applications

Our model may be able to shed light on a number of financial phenomena. We now discuss some of these potential applications. We divide the applications into those that relate to trading behavior (Section 5.1) and those that relate to asset prices (Section 5.2). In Section 5.3, we briefly discuss a few of the testable predictions that emerge from our framework.

5.1. Trading behavior

5.1.1. The disposition effect

The disposition effect is the finding that individual investors have a greater propensity to sell stocks that have gone *up* in value since purchase, rather than stocks that have gone down in value (Shefrin and Statman, 1985; Odean, 1998). This fact has turned out to be something of a puzzle, in that the most obvious potential explanations fail to capture important features of the data. Consider, for example, the most obvious potential explanation of all, the "informed trading" hypothesis. Under this view, investors sell stocks that have gone up in value because they have private information that these stocks will subsequently fall, and they hold on to stocks that have gone down in value because they have private information that these stocks will rebound. The difficulty with this view, as Odean (1998) points out, is that the prior winners people sell subsequently do *better*, on average, than the prior losers

---

condition $V_i(W, W) = 0$ is determined by buyer behavior, not seller behavior. To see this, suppose that an investor is trying to sell stock i. If $V_i(W, W) > 0$, then *all* investors holding the risk-free asset will want to switch to the stock and the market will fail to clear. On the other hand, if $V_i(W, W) < 0$, there will be no one for the seller to trade with: no one holding the risk-free asset will want to switch to the stock. Only if $V_i(W, W) = 0$ can we clear the market. The fact that prices are set by buyers has an important corollary: it means that the price of a stock does not depend on the average cost basis of the investors holding it.

they hold on to. Odean (1998) also considers other potential explanations based on taxes, rebalancing, and transaction costs, but argues that none of them is fully satisfactory.

Our analysis shows that a model that combines realization utility with a sufficiently positive time discount rate predicts a strong disposition effect. Unless forced to sell at a loss by a liquidity shock, the investor in our model *only* sells stocks trading at a price higher than the original purchase price.

In simple two-period settings, Shefrin and Statman (1985) and Barberis and Xiong (2009) show that realization utility, with no time discounting but with a functional form for utility that, as in prospect theory, is concave over gains and convex over losses, can predict a disposition effect. This paper proposes a related but distinct view of the disposition effect, namely that it arises from realization utility with a *linear* functional form for utility and a *positive* time discount rate.

We emphasize that realization utility does not, on its own, predict a disposition effect. In other words, to generate a disposition effect, it is not enough to assume that the investor derives pleasure from realizing a gain and pain from realizing a loss. We need an extra ingredient in order to explain why the investor would want to realize a gain *today*, rather than hold out for the chance of realizing an even bigger gain tomorrow. Shefrin and Statman (1985) and Barberis and Xiong (2009) point out one possible extra ingredient: a prospect theory functional form for utility. Such a functional form indeed explains why the investor would expedite realizing a gain and postpone realizing a loss. Here, we propose an alternative extra ingredient: a sufficiently positive time discount rate.

Our model is also well-suited for thinking about the disposition-type effects that have been uncovered in other settings. Genesove and Mayer (2001), for example, find that homeowners are reluctant to sell their houses at prices below the original purchase price. Our analysis shows that a model that combines linear realization utility with a positive time discount rate can capture this evidence.

Of all the applications we discuss in Section 5, the disposition effect is the most obvious, in the sense that it is very clear how the effect follows from our initial assumptions. However, as we noted in the Introduction, realization utility is in no sense a relabeling of the disposition effect. On the contrary, it is just one of a number of possible theories of the disposition effect, and can be distinguished from other theories through carefully constructed tests.

An example of a test that distinguishes various theories of the disposition effect can be found in Weber and Camerer (1998). These authors test the realization utility view of the disposition effect against the alternative view that it stems from an irrational belief in mean-

reversion. In a laboratory setting, they ask subjects to trade six stocks over a number of periods. In each period, each stock can either go up or down. The six stocks have different probabilities of going up in any period, ranging from 0.35 to 0.65, but subjects are not told which stock is associated with each possible up-move probability.

Weber and Camerer (1998) find that, just as in field data, their subjects exhibit a disposition effect. To try to understand the source of the effect, the authors consider an additional experimental condition in which the experimenter liquidates subjects' holdings and then tells them that they are free to reinvest the proceeds in any way they like. If subjects were holding on to their losing stocks because they thought that these stocks would rebound, we would expect them to re-establish their positions in these losing stocks. In fact, subjects do *not* re-establish these positions. This casts doubt on the mean-reversion view of the disposition effect and lends support to the realization utility view, namely that subjects were refusing to sell their losers simply because it would have been painful to do so. Under this view, subjects were relieved when the experimenter intervened and did it for them.[12]

5.1.2. Excessive trading

Using a database of trading activity at a large discount brokerage firm, Barber and Odean (2000) show that, after transaction costs, the average return of the individual investors in their sample falls below the returns on a range of benchmarks. This is puzzling: why do people trade so much if their trading hurts their performance? Barber and Odean (2000) consider a number of potential explanations, including taxes, rebalancing, and liquidity needs, but conclude that none of them can fully explain the patterns they observe.

Our model offers an explanation for this post-transaction-cost underperformance. Under this view, the investors in Barber and Odean's (2000) sample are guided by realization utility. This leads them to trade: specifically, to sell stocks that have risen in value since purchase so that they can enjoy bursts of positive utility, and to then invest the proceeds in new stocks. However, by trading, they incur transaction costs that cause them to underperform the benchmarks.

It is possible to compute the probability that the investor in our model sells a stock within any given interval of time after the initial purchase. Doing so will help us compare the trading frequency predicted by our model with that observed in actual brokerage accounts. When the investor first establishes a position in a stock, at time 0, say, we have $g_0 = 1$. When g_t reaches an upper barrier $g_* > 1$ or when a liquidity shock arrives, he sells the stock. To

---

[12]See Kaustia (2010) for additional evidence against the mean-reversion view of the disposition effect.

compute the probability that the investor sells the stock within s periods after establishing the position, we therefore need to compute the probability that g_t reaches g_* in the interval $(0, s)$ or that there is a liquidity shock during the same interval. The next proposition, which we prove in the Appendix, reports the result of this calculation.

Proposition 4. *The probability that the investor sells a stock within s periods of the date of purchase is*

$$
\begin{aligned}
G\left(s\right) \;\; = \;\; & 1 - e^{-\rho s} \\
& + e^{-\rho s}\left[N\left(\frac{-\ln g_* + \left(\mu - \frac{\sigma^2}{2}\right)s}{\sigma\sqrt{s}} \right) + e^{\left(\frac{2\mu}{\sigma^2} - 1\right)\ln g_*} N\left(\frac{-\ln g_* - \left(\mu - \frac{\sigma^2}{2}\right)s}{\sigma\sqrt{s}} \right) \right].
\end{aligned}
\tag{42}
$$

The expression in the square parentheses in (42) is the probability that g_t reaches g_* in the interval $(0, s)$. With this information in hand, it is easy to interpret the equation. The investor trades during the interval $(0, s)$ if one of two mutually exclusive events occurs: if there is a liquidity shock in $(0, s)$; or if there is no liquidity shock in $(0, s)$ but g_t reaches g_* in $(0, s)$. The probability of a trade in $(0, s)$ is therefore the probability of a liquidity shock in $(0, s)$, namely $1 - e^{-\rho s}$, plus the probability of no liquidity shock, namely $e^{-\rho s}$, multiplied by the probability that g_t reaches g_*.

Fig. 5 shows how the probability of selling a stock within a year of purchase, $G(1)$, depends on the model parameters. To construct the graphs, we use the model of Section 3 which allows for a transaction cost, a liquidity shock, and piecewise-linear utility. For any given parameter values, we compute the liquidation point g_* from (22)–(25) and substitute the result into the expression for $G(1)$ in Proposition 4. The graphs vary each of μ, σ, δ', k, and λ in turn, keeping the remaining parameters fixed at their benchmark values

$$
\begin{aligned}
(\alpha, \mu, \sigma, k) \;\; &= \;\; (0.015, 0.015, 0.5, 0.005) \\
(\delta', \rho, \lambda, \beta) \;\; &= \;\; (0.08, 0.1, 1.5, 1).
\end{aligned}
\tag{43}
$$

Some of the results in Fig. 5 are not very surprising. The middle-left graph shows that as the investor becomes more impatient, the probability of a trade rises. And the middle-right graph shows that as the transaction cost falls, the probability of a trade again rises.

The top-left and top-right graphs, which vary μ and σ respectively, are less predictable. In both cases, there are two factors at work. On the one hand, for any fixed liquidation point g_*, a higher μ or σ raises the likelihood that g_* will be reached within the year-long interval. However, as we saw in Fig. 2, the liquidation point g_* itself goes up as μ and σ go

25

up, thereby *lowering* the chance that g_* will be reached. Without computing $G(1)$ explicitly, it is hard to know which factor will dominate.

The top graphs in Fig. 5 show that, interestingly, a different factor dominates in each case. As μ rises, the probability of a trade falls. Roughly speaking, as μ rises, the liquidation point rises more quickly than the stock's ability to reach it. As σ rises, however, the probability of a trade goes *up*: in this case, the liquidation point rises less quickly than the stock's ability to reach it.

The bottom-left graph, which varies λ, shows that the probability of a trade declines as the sensitivity to losses rises. If λ is high, the investor is reluctant to sell a stock trading at a gain because if he does, he will have to buy a new stock, which might go down and which he might be forced to sell at a loss by a liquidity shock.

Barber and Odean (2000) find that in their sample of households with brokerage accounts, the mean and median annual turnover rates are 75% and 30%, respectively. Fig. 5 shows that for the benchmark parameter values, our model predicts a trading frequency that is of a similar order of magnitude. When $\sigma = 0.5$, for example, the probability that an investor trades a specific stock in his portfolio within a year of purchase is approximately 0.6. Of course, the fact that the trading frequency predicted by our model is similar to that observed in actual brokerage accounts is not an accident: we *chose* the benchmark value of δ' to ensure that this would be the case.

When we say that realization utility can help us understand "excessive trading," we do not mean that it can explain the high overall volume of trading in financial markets. Rather, we mean something narrower: that it can help us understand why individual investors trade as much as they do in their brokerage accounts, given that they would earn higher returns, on average, if they traded less. While realization utility investors are keen to trade a stock that has risen in value, they are *not* keen to trade a stock that has fallen in value. It is therefore an open question as to whether an increase in the fraction of investors in the economy who are guided by realization utility would lead to an increase in the overall volume of trading.

5.1.3. Underperformance before transaction costs

Some studies find that the average individual investor underperforms benchmarks even before transaction costs (Barber et al., 2009). Our model may be able to shed light on this by way of one of the predictions we discussed in Sections 2 and 3: that an investor who thinks in terms of realization utility is often willing to buy a stock with a low expected return, so long as the stock's volatility is sufficiently high.

Suppose that the investing population consists of two groups: individuals, who think in

terms of realization utility; and institutions, who do not. Since individuals are guided by realization utility, they may be more willing than institutions to buy stocks with low expected returns. Moreover, since the average portfolio return before transaction costs across *all* investors must equal the market return, we should observe the average individual underperforming market benchmarks before transaction costs and the average institution outperforming the benchmarks, again before transaction costs. This prediction is broadly consistent with the available evidence.[13]

5.1.4. Trading volume in rising and falling markets

Researchers have found that in many different asset classes, trading volume is higher in rising markets than in falling markets (Stein, 1995; Statman, Thorley, and Vorkink, 2006; Griffin, Nardari, and Stulz, 2007). Robust though this finding is, there are few explanations for it. The equilibrium model of Section 4 offers a way of understanding it. In that model, there is indeed more trading in rising markets. In a rising market, the stocks held by realization utility investors start hitting their liquidation points. When this happens, these investors sell their stocks to other realization utility investors. As a result, trading volume goes up.

The same line of reasoning can motivate the use of turnover as a measure of investor sentiment (Baker and Wurgler, 2007). If some investors have very positive sentiment and push stock prices up as a result, realization utility investors will start trading heavily. This creates a link between turnover and sentiment.

5.1.5. The effect of historical highs on the propensity to sell

Our model implies that there will be more trading in rising markets, but it can also make more precise predictions as to how trading activity will vary over time. For example, it predicts that individual investors—the investor group that is more likely to think in terms of realization utility—will have a much higher propensity to sell a stock once its price moves above its historical high.

---

[13]So far, our model has pointed to two ways in which realization utility can lower an investor's Sharpe ratio: it leads him to buy stocks with low expected returns and high volatility; and by encouraging him to trade, it leads him to incur transaction costs. There is one more channel through which realization utility can harm the investor's performance—a channel that, while important, lies outside our model. A strategy that sells winners but holds on to losers will lower the investor's average return if his typical holding period coincides with the horizon at which stocks exhibit momentum. At least for some investors, this does appear to be the case: the investors in Barber and Odean's (2000) sample hold stocks for a few months, on average—a horizon at which stock returns exhibit significant momentum.

To see this, consider a stock that, on January 1st, is trading at $30. Suppose that it then rises through January and February, reaching a high of $45 by February 28th. It then declines significantly through most of March but, towards the end of March, starts rising again, passing through the previous high of $45 on March 31st and continuing upwards.

Our model predicts that after the stock passes $45 on March 31st, there will be a sharp increase in selling by individual investors. To see why, note that there will be very little selling between February 28th and March 31st. During this time, the stock is trading below its high of $45. The only investors who would want to sell in this interval are those targeting liquidation points below $45. But the majority of these investors *will have sold the stock already*, before February 28th, when the stock first reached $45. Once the stock moves above $45 on March 31st, however, investors targeting liquidation points higher than $45 will start selling. As claimed above, then, individual investors' propensity to sell a stock will increase sharply as the stock price moves above its historical high.

Our prediction is consistent with the available evidence. Grinblatt and Keloharju (2001) find that households' propensity to sell a stock *does* increase strongly once the stock price moves above its historical high for that month. Similarly, albeit in a different context, Heath, Huddart, and Lang (1999) find that executives are much more likely to exercise stock options when the underlying stock price exceeds its historical high. Finally, Baker, Pan, and Wurgler (2009) show that, when a firm makes a takeover bid for another firm, the offer price is more likely to slightly exceed the target's 52-week historical high than to be slightly below it; and that there is a discontinuous increase in deal success as the offer price rises through the 52-week high. This is consistent with the idea that, as a consequence of realization utility, investors are more likely to sell their shares in the target company at a price that exceeds the historical high.[14]

5.1.6. The individual investor preference for volatile stocks

Kumar (2009) analyzes the trades of approximately 60,000 households with accounts at a large discount brokerage firm. He finds that, as a group, the individual investors in his sample overweight highly volatile stocks: these stocks make up a larger fraction of the value

---

[14]It is tempting to interpret Grinblatt and Keloharju's (2001) finding as evidence that investors use the historical high as an *explicit* reference point: for example, that they derive utility from the difference between the price at which they sell a stock and its historical high. Our analysis shows, however, that Grinblatt and Keloharju's (2001) result can arise in a model in which the only explicit reference point is the purchase price. The historical high emerges as a reference point endogeneously because of the nature of the investor's optimal strategy.

of the aggregate individual investor portfolio, constructed using these data, than they do of the aggregate market portfolio. Realization utility offers a way of understanding this. As we saw in Sections 2 and 3, investors who are guided by realization utility often have a strong preference for volatile stocks. Moreover, these investors are more likely to be individuals than institutions.

5.2. Asset pricing

Our model may also be helpful for understanding certain asset pricing patterns. We now discuss three applications of this type.

5.2.1. The low average return of volatile stocks

Ang et al. (2006) show that, in the cross-section, and after controlling for previously known predictor variables, a stock's daily return volatility over the previous month *negatively* predicts its return in the following month. This finding, which holds not only in the U.S. stock market but in many international stock markets as well, is puzzling. Even if we allow ourselves to think of a stock's own volatility as risk, the result is the opposite of what we would expect: it says that "riskier" stocks have *lower* average returns.

Our model offers a novel explanation for this finding. We noted earlier—see the middle-right graph in Fig. 2—that for some parameter values, realization utility investors are risk seeking. As a result, they will exert heavy buying pressure on stocks that are highly volatile. These stocks may then become overpriced. If so, their subsequent average return will indeed be low.

We now check this intuition using the equilibrium model of Section 4. We assign all investors the same benchmark parameter values

$$(\delta', \rho, \lambda, \beta) = (0.08, 0.1, 1.5, 1), \tag{44}$$

and assume that the excess dividend growth rate and the transaction cost are the same for all stocks, namely $\mu = -0.03$ and $k = 0.005$, respectively. For values of σ ranging from 0.01 to 0.5, we use equilibrium condition (41) to compute the dividend yield α and hence the expected excess return $\alpha + \mu$ that a stock with any given standard deviation must earn in order for its market to clear.[15]

---

[15]Since μ is the *excess* dividend growth rate, a negative value of μ does not necessarily mean that the dividend growth rate is negative, just that it is below the risk-free rate. Since, for the parameter values in (44), the investors in our economy are risk seeking, the dividend growth rate *must* be below the risk-free rate to prevent prices from exploding, just as, in a standard Gordon growth model with risk-neutral investors,

The top-left graph in Fig. 6 plots the resulting relationship between standard deviation and expected excess return. The graph confirms our prediction: more volatile stocks earn lower average returns; in this sense, they are overpriced.[16]

The top-left graph also shows that for the parameter values in (44), stocks earn negative average excess returns, which is inconsistent with the positive historical equity premium. A negative equity premium is not a generic prediction of our model: for values of ρ and λ that are somewhat higher than those in (44), and for values of δ' that are somewhat lower, the investors become risk averse rather than risk seeking and the equity premium turns positive. It is difficult, however, for the homogeneous agent economy we are analyzing to generate both a positive equity premium *and* a negative relationship between volatility and average return in the cross-section. We conjecture that it may be possible to generate both of these facts in an economy with heterogeneous realization utility investors, some of whom are risk seeking and some of whom are risk averse.

Another way of reconciling the top-left graph with the positive historical equity premium is to say that the result in the graph only applies to stocks that are primarily held by investors who think in terms of realization utility—most likely, individual investors. Since these stocks constitute a small fraction of the total stock market capitalization, they play only a minor role in determining the aggregate equity premium. One prediction of this view is that the cross-sectional relationship between volatility and average return documented by Ang et al. (2006) should be stronger among stocks traded by individual investors. This is exactly the finding of Han and Kumar (2011).

5.2.2. The heavy trading of highly valued assets

A robust empirical finding is that assets that are highly valued, and possibly overvalued, are also heavily traded (Hong and Stein, 2007). Growth stocks, for example, are more heavily traded than value stocks; the highly priced technology stocks of the late 1990s changed hands at a rapid pace; and shares at the center of famous bubble episodes, such as those of the East India Company at the time of the South Sea bubble, also experienced heavy trading.

Our model may be able to explain this coincidence of high prices and heavy trading.

the dividend growth rate has to be below the risk-free rate. Note that a negative excess dividend growth rate μ does not necessarily imply a negative expected excess return. The expected excess return is $\alpha + \mu$. This can be positive even if μ is negative.

[16]In our model, the risky assets are infinitely lived. We have studied a variant of the model in which the risky assets stochastically "expire" based on the arrival of Poisson-distributed liquidation shocks. We find that in an economy with realization utility investors, a short-horizon asset—one with a higher liquidation shock intensity—can earn a higher Sharpe ratio than a long-horizon asset.

Specifically, it predicts that this phenomenon will occur for assets whose value is especially uncertain.

Suppose that the uncertainty about an asset's value goes up, thereby increasing σ, the standard deviation of returns. As noted earlier, investors who think in terms of realization utility will now find the asset more attractive. If there are many such investors in the economy, the asset's price will be pushed up.

At the same time, the top-right graph in Fig. 5 shows that as σ goes up, the probability that an investor will trade the asset also goes up: simply put, a more volatile asset tends to reach its liquidation point more rapidly. In this sense, the overvaluation will coincide with higher turnover, and this will occur when uncertainty about the asset's value is especially high. Under this view, the late 1990s were years when realization utility investors, attracted by the high uncertainty of technology stocks, bought these stocks, pushing their prices up; as (some of) these stocks rapidly reached their liquidation points, the realization utility investors sold them and then immediately bought new ones.

We now check this intuition using the equilibrium framework of Section 4. As in our discussion of the low average return of volatile stocks, we assign all investors the benchmark parameter values in (44) and assume that the excess dividend growth rate and the transaction cost are the same for all stocks, namely $\mu = -0.03$ and $k = 0.005$, respectively. For values of σ ranging from 0.01 to 0.5, we again use condition (41) to compute the corresponding equilibrium expected excess return; but this time, as a guide to the intensity of trading, we also use (42) to compute $G(1)$, the probability of a trade within a year of purchase.

The top-right graph in Fig. 6 plots the resulting relationship between the expected excess return and the trade probability. It confirms that stocks with lower expected returns—stocks that are more "overpriced"—do indeed experience more turnover.

5.2.3. Momentum

Grinblatt and Han (2005) study an economy in which some investors' demand for a stock depends, negatively, on the difference between the current stock price and the price they paid for the stock. They show that in this economy, as in actual data, stock returns exhibit momentum. The authors suggest one possible foundation for the demand function they propose, namely, a combination of prospect theory and mental accounting. Our model suggests a different, albeit related foundation: linear realization utility. In combination with a sufficiently positive time discount rate, linear realization utility also leads to a demand function for a stock that depends, negatively, on the difference between the current stock price and the purchase price. This, in turn, suggests that momentum may ultimately stem,

31

at least in part, from realization utility.

A limitation of the pricing model in Section 4 is that it does not allow us to illustrate the link between realization utility and momentum: in that model, stock returns are not predictable. To see why the link breaks down, recall the original intuition for it. The idea is that if a stock rises in value, realization utility investors will start selling it in order to realize a gain. This selling pressure causes the stock to become undervalued. Subsequently, the stock price moves higher, on average, as it corrects from this undervalued point to a more reasonable valuation. An upward price move is therefore followed by another upward price move, on average. This generates a momentum effect in the cross-section of stock returns.

In our model, realization utility investors do indeed start selling when a stock rises in value. However, this does *not* depress the stock price because of the perfectly elastic demand for the stock from other realization utility investors. As a result, there is no momentum. We suspect that the link between realization utility and momentum can be formalized in an economy with both realization utility investors *and* expected utility investors. In such an economy, when realization utility investors sell a stock that is rising in value, their selling *will* depress the stock price because the demand from expected utility investors will not be perfectly elastic.

5.3. Testable predictions

In Sections 5.1 and 5.2, we argue that realization utility offers a simple way of understanding a range of financial phenomena. In this section, we briefly note a few of the new predictions that emerge from our framework.

One set of predictions is based on the graphs in Fig. 5, which show how the probability of trade depends on various parameters. One of these predictions, that the investor is more likely to trade a stock within a year of purchase when transaction costs are lower, is not unique to our model. However, the figure also suggests some other, more novel predictions: that the probability that the investor trades a stock within a year of purchase is an increasing function of his impatience and of the stock's volatility, and a decreasing function of his sensitivity to losses.

The prediction relating the probability of trade to a stock's volatility is straightforward to test empirically. To test the predicted link between trade probability and impatience and between trade probability and sensitivity to losses, we need estimates of impatience and loss sensitivity, which may be difficult to obtain. In recent years, however, researchers have pioneered clever techniques for extracting information about investors' psychological profiles. Grinblatt and Keloharju (2009), for example, use military test scores from Finland

to estimate overconfidence. This success makes us more optimistic that a test of the link between trade probability on the one hand, and impatience and loss sensitivity on the other, can also be implemented.

If we are indeed able to measure investor impatience, there are other predictions that can be tested. As noted earlier, two of the more striking implications of realization utility—that investors will be willing to buy stocks that are highly volatile and that have low expected returns—depend crucially on the discount rate δ. Roughly speaking, a stock with a low expected return or with high volatility offers the investor the prospect of realizing either a short-term gain or a long-term loss. The higher the discount rate δ, the more attractive this tradeoff becomes. In short, then, if we are able to measure investor impatience, we should find that more impatient investors allocate more to stocks with low expected returns, thereby earning low portfolio returns even before taking transaction costs into account; and also that they tilt their portfolios more heavily towards volatile stocks.

6. Conclusion

A number of authors have suggested that investors may derive utility from realizing gains and losses. We present a model of this "realization utility," study its predictions, and show that it can shed light on a number of puzzling facts.

There are several possible directions for future research. First, while many of our model's implications match the observed facts, some do not. For example, our model predicts too strong a disposition effect: in our framework, investors never voluntarily sell stocks at a loss, while, in reality, they clearly do. It would be useful to see whether an extension of our model—one that modifies our preference specification in some way, or that allows for richer beliefs about expected stock returns—can make more accurate predictions.[17]

Another natural research direction involves testing the implications of realization utility. To do this, we can use field data on investor trading behavior; or experimental data, as in Weber and Camerer (1998). Another type of data that has recently become available is neural data. For example, Frydman et al. (2011) use functional magnetic resonance imaging (fMRI) technology to monitor the brain activity of 28 subjects while they trade stocks in an experimental market. The authors use the neural data to test some theories of investor behavior, including the one presented in this paper.

Finally, it would be useful to think about other applications of realization utility. These applications may again concern the trading and pricing of financial securities, or they may be

---

[17]Two recent studies that take up this question are Ingersoll and Jin (2011) and Henderson (2012).

drawn from quite different areas of study. After all, the core idea that, in our view, underlies realization utility—that people break their experiences down into episodes and receive a burst of utility when an episode comes to an end—strikes us as one that may be relevant in many contexts, not just the financial market context that we have focused on in this paper.

Appendix

Proof of Proposition 1. At time t, the investor can either liquidate his position or hold it for an infinitesimal period dt. We therefore have

$$
\begin{aligned}
&V\left(W_t, B_t\right)\\
&= \max\left\{u\left((1-k)W_t - B_t\right) + V\left((1-k)W_t, (1-k)W_t\right),\right. \hspace{2cm} (45)\\
&\qquad \left. v(D_{i,t})dt + (1-\rho dt)E_t[e^{-\delta dt}V\left(W_{t+dt}, B_{t+dt}\right)] + \rho dt\left[u\left((1-k)W_t - B_t\right)\right]\right\}.
\end{aligned}
$$

The first argument of the "max" function corresponds to the case where the investor liquidates his position at time t: he receives realization utility of $u((1-k)W_t - B_t)$ and cash proceeds of $(1-k)W_t$ which he immediately invests in another stock. The second argument of the "max" function corresponds to the case where the investor instead holds his position for an infinitesimal period dt: he receives utility $v(D_{i,t})dt$ from the flow of dividends; with probability $e^{-\rho dt} \approx 1 - \rho dt$, there is no liquidity shock during the interval and his value function is the expected future value function discounted back; and with probability $1 - e^{-\rho dt} \approx \rho dt$, there *is* a liquidity shock, in which case he sells his holdings, exits the asset markets, and receives realization utility of $u((1-k)W_t - B_t)$.

Given the homogeneity property in (8), we can write the value function as

$$
V\left(W_t, B_t\right) = B_t U\left(g_t\right).
$$

Substituting this into (45), cancelling B_t from both sides, and applying Ito's lemma gives

$$
\begin{aligned}
U\left(g_t\right) &= \max\{u((1-k)g_t - 1) + (1-k)g_t U\left(1\right), \hspace{3cm} (46)\\
&\quad U(g_t) + \left[\alpha\beta g_t + \frac{1}{2}\sigma^2 g_t^2 U''\left(g_t\right) + \mu g_t U'\left(g_t\right) - (\rho + \delta')U\left(g_t\right) + \rho u\left((1-k)g_t - 1\right)\right]dt\}.
\end{aligned}
$$

Eq. (46) implies that any solution to (10) must satisfy

$$
U\left(g_t\right) \geq u((1-k)g_t - 1) + (1-k)g_t U\left(1\right) \hspace{3cm} (47)
$$

and

$$
\alpha\beta g_t + \frac{1}{2}\sigma^2 g_t^2 U''\left(g_t\right) + \mu g_t U'\left(g_t\right) - (\rho + \delta')U\left(g_t\right) + \rho u\left((1-k)g_t - 1\right) \leq 0. \hspace{1cm} (48)
$$

Formally speaking, the decision problem in (10) is an optimal stopping problem. To solve it, we first construct a function $U(g_t)$ that satisfies conditions (47) and (48) and that is both continuous and continuously differentiable—this last condition is sometimes known as the "smooth pasting" condition. If we are able to do this, then, given that certain technical conditions are satisfied, the constructed function $U(g_t)$ will indeed be a solution to problem (10).

We construct $U(g_t)$ in the following way. If g_t is low, specifically, if $g_t \in (0, g_*)$, we suppose that the investor continues to hold his current position. In this "continuation" region, condition (48) holds with equality. If g_t is sufficiently high, specifically, if $g_t \in (g_*, \infty)$, we suppose that the investor liquidates his position. In this "liquidation" region, condition (47) holds with equality. As in the statement of the proposition, we refer to g_* as the liquidation point.

Since $u(\cdot)$ is linear, the value function $U(\cdot)$ in the continuation region satisfies

$$\frac{1}{2}\sigma^2 g_t^2 U''(g_t) + \mu g_t U'(g_t) - (\rho + \delta') U(g_t) + (\alpha\beta + \rho(1-k))g_t - \rho = 0.$$

The solution to this equation is

$$U(g_t) = a g_t^{\gamma_1} + \frac{\alpha\beta + \rho(1-k)}{\rho + \delta' - \mu} g_t - \frac{\rho}{\rho + \delta'} \quad \text{for} \quad g_t \in (0, g_*), \tag{49}$$

where γ_1 is given in (15) and where a is determined below.

In the liquidation region, we have

$$U(g_t) = (1-k)g_t(1 + U(1)) - 1. \tag{50}$$

Note that the liquidation point g_* satisfies $g_* \geq 1$. For if $g_* < 1$, then $g_t = 1$ would fall into the liquidation region, which, from (50), would imply

$$U(1) = (1-k)U(1) - k.$$

For $k > 0$ and $U(1) \geq 0$, this is a contradiction. Since $g_* \geq 1$, then, we infer from (49) that

$$U(1) = a + \frac{\alpha\beta + \rho(1-k)}{\rho + \delta' - \mu} - \frac{\rho}{\rho + \delta'}. \tag{51}$$

The value function must be continuous and continuously differentiable at the liquidation point g_*. This implies

$$a g_*^{\gamma_1} + \frac{\alpha\beta + \rho(1-k)}{\rho + \delta' - \mu} g_* - \frac{\rho}{\rho + \delta'} = (1-k)g_*(1 + U(1)) - 1$$

$$a\gamma_1 g_*^{\gamma_1 - 1} + \frac{\alpha\beta + \rho(1-k)}{\rho + \delta' - \mu} = (1-k)(1 + U(1)).$$

35

Solving these two equations, we obtain the expression for a in (16) and the nonlinear equation for g_* in (17). It is straightforward to check that if restriction (11) holds, Eq. (17) has a unique solution in the range $(1, \infty)$.

We now verify that the function $U(g_t)$ summarized in Eq. (14) satisfies conditions (47) and (48). Define

$$f(g) \equiv (1-k)(1+U(1))g - 1.$$

By construction, $f(g)$ is a straight line that coincides with $U(g)$ for $g \geq g_*$. Since $\gamma_1 > 1$—this follows from $\mu < \rho + \delta'$ which, in turn, follows from restriction (11)—$U(g)$ in Eq. (14) is a convex function. It must therefore lie above the straight line $f(g)$ for all $g < g_*$. Condition (47) is therefore satisfied.

We now check that condition (48) holds. Define

$$H(g) \equiv \frac{1}{2}\sigma^2 g^2 U''(g) + \mu g U'(g) - (\rho + \delta') U(g) + (\alpha\beta + \rho(1-k))g - \rho.$$

For $g < g_*$, $H(g) = 0$ by construction. For $g \geq g_*$, $U(g) = f(g)$, so that

$$H(g) = -(1-k)g\left[(\rho + \delta' - \mu)(1+U(1)) - (\rho + \frac{\alpha\beta}{1-k})\right] + \delta'.$$

Substituting (51) and (16) into this expression, we obtain

$$
\begin{aligned}
H(g) &= -(1-k)g\left\{\frac{\delta'(\rho+\delta'-\mu)}{\rho+\delta'}\left[1 + \frac{1}{(\gamma_1-1)g_*^{\gamma_1}}\right] - \frac{k}{1-k}(\alpha\beta + \rho(1-k)) - \frac{\delta'}{(1-k)g}\right\} \\
&\leq -(1-k)g\left\{\frac{\delta'(\rho+\delta'-\mu)}{\rho+\delta'}\left[1 + \frac{1}{(\gamma_1-1)g_*^{\gamma_1}}\right] - \frac{k}{1-k}(\alpha\beta + \rho(1-k)) - \frac{\delta'}{(1-k)g_*}\right\} \\
&= -\frac{g}{g_*}\frac{\delta'}{(\rho+\delta')(\gamma_1-1)}(\rho+\delta'-\mu\gamma_1).
\end{aligned}
$$

The last equality follows by applying (17). Using (15), it is straightforward to show that if $\mu < \rho + \delta'$, as assumed in restriction (11), then $\rho + \delta' - \mu\gamma_1 > 0$. Therefore, $H(g) < 0$ for $g \geq g_*$, thereby confirming that condition (48) holds for all $g_t \in (0, \infty)$.

To formally complete the derivation of Proposition 1, we have proved a verification theorem. This theorem uses the fact that conditions (47) and (48) hold everywhere to confirm that the stopping strategy proposed above is indeed the optimal one. For space reasons, we do not present the details of this step here.

Proof of Proposition 2. The proof is very similar in structure to the proof of Proposition 1. We therefore present only the key steps. From (8), the value function takes the form

$$V(W_t, B_t) = B_t U(g_t).$$

36

Following the same reasoning as in the proof of Proposition 1, we find that $U(\cdot)$ again satisfies Eq. (46) and inequalities (47) and (48). The only difference is that $u(\cdot)$ now has the piecewise-linear form in (18).

As before, we conjecture two regions: a continuation region, $g_t \in (0, g_*)$, and a liquidation region, $g_t \in (g_*, \infty)$. In the continuation region, $U(\cdot)$ satisfies

$$\frac{1}{2}\sigma^2 g_t^2 U''(g_t) + \mu g_t U'(g_t) - (\rho + \delta')U(g_t) + \alpha\beta g_t + \rho u((1-k)g_t - 1) = 0. \qquad (52)$$

The form of the $u(\cdot)$ term depends on whether its argument, $(1-k)g_t - 1$, is greater or less than zero. Note that the cross-over point, $g_t = \frac{1}{1-k}$, lies below g_*, so that $g_* \geq \frac{1}{1-k}$. For if $g_* < \frac{1}{1-k}$, then $g_t = \frac{1}{1-k}$ would be in the liquidation region, which, from (20), would imply

$$U\left(\frac{1}{1-k}\right) = U(1),$$

contradicting the reasonable restriction that $U(g_t)$ be strictly increasing in g_t. Since $g_* \geq \frac{1}{1-k}$, we further subdivide the continuation region $(0, g_*)$ into two subregions, $\left(0, \frac{1}{1-k}\right)$ and $\left(\frac{1}{1-k}, g_*\right)$.

For $g_t \in \left(0, \frac{1}{1-k}\right)$, (52) becomes

$$\frac{1}{2}\sigma^2 g_t^2 U''(g_t) + \mu g_t U'(g_t) - (\rho + \delta')U(g_t) + (\alpha\beta + \rho\lambda(1-k))g_t - \rho\lambda = 0.$$

The solution to this equation is

$$U(g_t) = bg_t^{\gamma_1} + \frac{\alpha\beta + \rho\lambda(1-k)}{\rho + \delta' - \mu}g_t - \frac{\rho\lambda}{\rho + \delta'} \quad \text{for} \quad g_t \in \left(0, \frac{1}{1-k}\right), \qquad (53)$$

where γ_1 is defined in (15) and where b is determined below.

For $g_t \in \left(\frac{1}{1-k}, g_*\right)$, (52) becomes

$$\frac{1}{2}\sigma^2 g_t^2 U''(g_t) + \mu g_t U'(g_t) - (\rho + \delta')U(g_t) + (\alpha\beta + \rho(1-k))g_t - \rho = 0.$$

The solution to this equation is

$$U(g_t) = c_1 g_t^{\gamma_1} + c_2 g_t^{\gamma_2} + \frac{\alpha\beta + \rho(1-k)}{\rho + \delta' - \mu}g_t - \frac{\rho}{\rho + \delta'} \quad \text{for} \quad g_t \in \left(\frac{1}{1-k}, g_*\right),$$

where

$$\gamma_2 = -\frac{1}{\sigma^2}\left[\sqrt{\left(\mu - \frac{1}{2}\sigma^2\right)^2 + 2(\rho + \delta')\sigma^2} + \left(\mu - \frac{1}{2}\sigma^2\right)\right] < 0$$

and where c_1 and c_2 are determined below.

The value function must be continuous and continuously differentiable at $g_t = \frac{1}{1-k}$. We therefore have

$$b\left(\frac{1}{1-k}\right)^{\gamma_1} = c_1\left(\frac{1}{1-k}\right)^{\gamma_1} + c_2\left(\frac{1}{1-k}\right)^{\gamma_2} - \frac{(\lambda-1)\mu\rho}{(\rho+\delta'-\mu)(\rho+\delta')},$$

which is (24), and

$$b\gamma_1\left(\frac{1}{1-k}\right)^{\gamma_1-1} = c_1\gamma_1\left(\frac{1}{1-k}\right)^{\gamma_1-1} + c_2\gamma_2\left(\frac{1}{1-k}\right)^{\gamma_2-1} - \frac{(\lambda-1)(1-k)\rho}{\rho+\delta'-\mu}.$$

Together, these equations imply Eq. (22).

In the liquidation region, $g_t \in (g_*, \infty)$, using the fact that $g_* \geq 1$, we have

$$U(g_t) = (1-k)g_t(1+U(1)) - 1.$$

The value function must be continuous and continuously differentiable at the liquidation point, so that

$$c_1 g_*^{\gamma_1} + c_2 g_*^{\gamma_2} + \frac{\alpha\beta + \rho(1-k)}{\rho+\delta'-\mu}g_* = (1-k)g_*(1+U(1)) - \frac{\delta'}{\rho+\delta'}$$

$$c_1\gamma_1 g_*^{\gamma_1-1} + c_2\gamma_2 g_*^{\gamma_2-1} + \frac{\alpha\beta + \rho(1-k)}{\rho+\delta'-\mu} = (1-k)(1+U(1)).$$

Since, from (53),

$$U(1) = b + \frac{\alpha\beta}{\rho+\delta'-\mu} + \frac{\rho\lambda(\mu-k\rho-k\delta')}{(\rho+\delta')(\rho+\delta'-\mu)},$$

we obtain

$$c_1 g_*^{\gamma_1} + c_2 g_*^{\gamma_2} + \frac{\alpha\beta + (1-k)(\mu-\delta')}{\rho+\delta'-\mu}g_* + \frac{\delta'}{\rho+\delta'} = (1-k)g_*\left(b + \frac{\alpha\beta}{\rho+\delta'-\mu} + \frac{\rho\lambda(\mu-k\rho-k\delta')}{(\rho+\delta')(\rho+\delta'-\mu)}\right),$$

which reduces to Eq. (25); and also Eq. (23).

Proof of Proposition 3. We solve the decision problem in (32) using the same technique as the one employed in the proofs of Propositions 1 and 2. In particular, we replace α, μ, σ, and k in (46) with α_i, μ_i, σ_i, and k_i—the dividend yield, expected excess capital gain, standard deviation, and transaction cost of stock i, respectively. We also note that $U_i(1) = 0$ in equilibrium. It is then straightforward to obtain the results in Proposition 3.

Proof of Proposition 4. Define $x_t \equiv \ln g_t$ and $x_* \equiv \ln g_*$. Then,

$$dx_t = \mu_x dt + \sigma dZ_t, \qquad \mu_x = \mu - \frac{\sigma^2}{2}.$$

If the investor has not yet traded, what is the probability that he trades at least once in the following s periods? Note that he will trade if the stock price rises sufficiently high so that the process x_t hits the barrier x_*; or if there is a liquidity shock. The probability is therefore a function of x_t and of the length of the period s. We denote it by $p(x, s)$.

Since a probability process is a martingale, its drift is zero, so that

$$-p_s + \mu_x p_x + \frac{1}{2} \sigma^2 p_{xx} + \rho(1 - p) = 0.$$

The last term on the left-hand side is generated by the liquidity shock: if a liquidity shock arrives, the probability of a trade jumps from p to 1. The probability function must also satisfy two boundary conditions. First, if the process x_t is already at the barrier x_*, there is a trade for sure:

$$p(x_*, s) = 1, \qquad \forall s \geq 0.$$

Second, if the length of the remaining time period is zero and the price level is such that $x < x_*$, there will be no trade:

$$p(x, 0) = 0, \qquad \forall x < x_*.$$

The solution to the differential equation, subject to the boundary conditions, is

$$p(x, s) = 1 - e^{-\rho s} + e^{-\rho s}\left[N\left(\frac{x - x_* + \mu_x s}{\sigma \sqrt{s}}\right) + e^{-\frac{2\mu_x}{\sigma^2}(x - x_*)} N\left(\frac{x - x_* - \mu_x s}{\sigma \sqrt{s}}\right)\right].$$

Substituting $x = 0$, $x_* = \ln g_*$, and $\mu_x = \mu - \frac{\sigma^2}{2}$ into this expression, we obtain the result in Proposition 4.
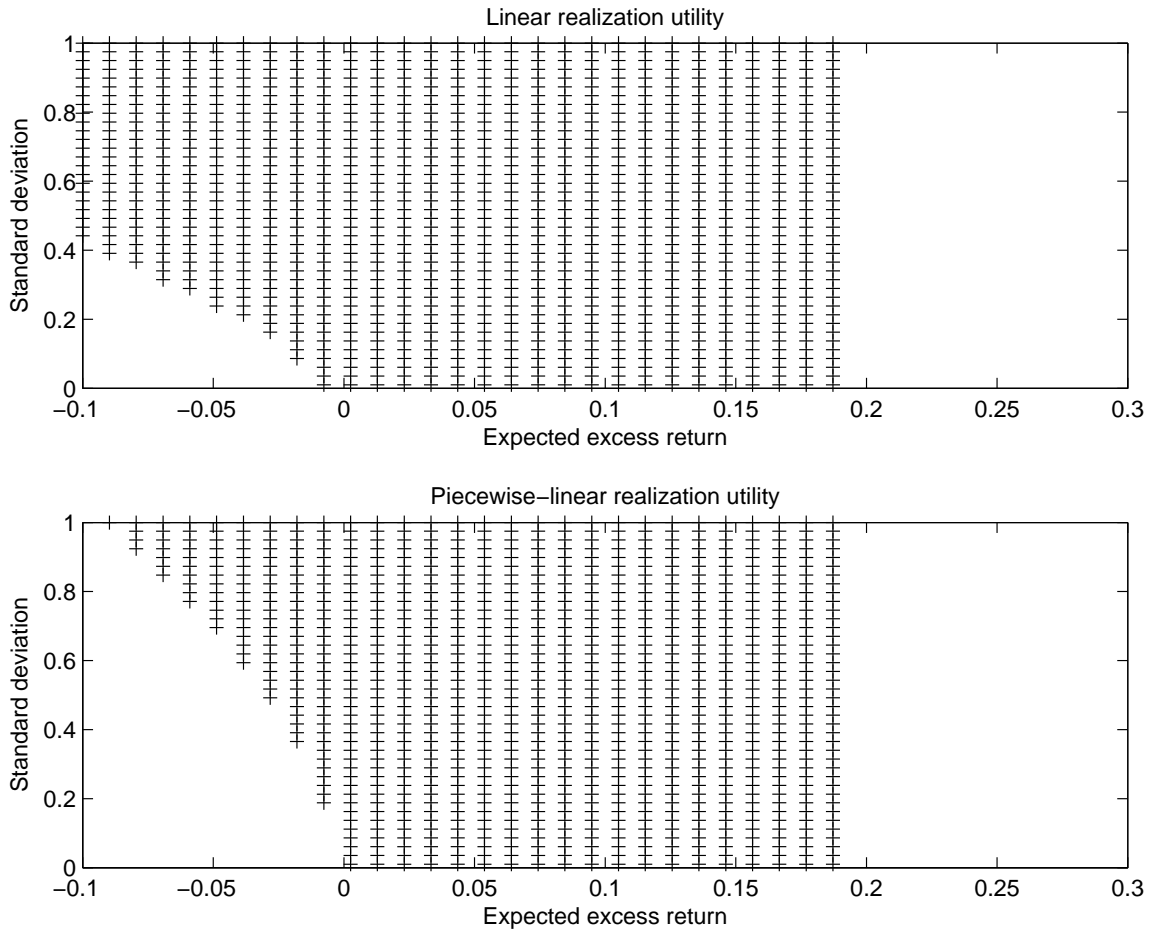
References

Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. Journal of Finance 61, 259–299.

Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. Journal of Economic Perspectives 21, 129–151.

Baker, M., Pan, X., Wurgler, J., 2009. A reference point theory of mergers and acquisitions. Unpublished working paper. Harvard University and New York University.

Barber, B., Odean, T., 2000. Trading is hazardous to your wealth: the common stock performance of individual investors. Journal of Finance 55, 773–806.

Barber, B., Lee, Y., Liu, Y., Odean, T., 2009. Just how much do investors lose from trade? Review of Financial Studies 22, 151–186.
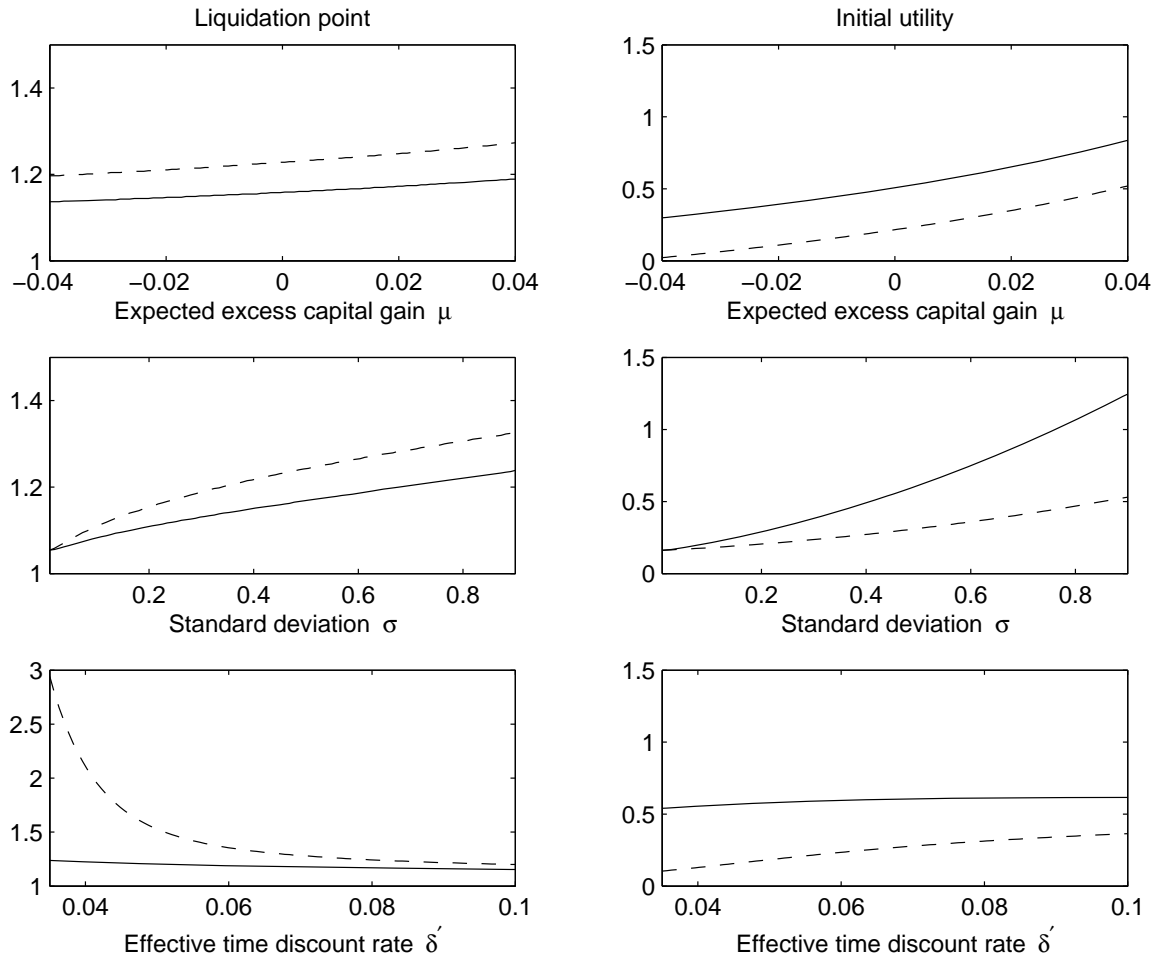
Barberis, N., Huang, M., 2001. Mental accounting, loss aversion, and individual stock returns. Journal of Finance 56, 1247–1292.

Barberis, N., Huang, M., Santos, T., 2001. Prospect theory and asset prices. Quarterly Journal of Economics 116, 1–53.

Barberis, N., Xiong, W., 2009. What drives the disposition effect? An analysis of a long-standing preference-based explanation. Journal of Finance 64, 751–784.

Benartzi, S., Thaler, R., 1995. Myopic loss aversion and the equity premium puzzle. Quarterly Journal of Economics 110, 73–92.

Frydman, C., Barberis, N., Camerer, C., Bossaerts, P., Rangel, A., 2011. Testing theories of investor behavior using neural data. Unpublished working paper. California Institute of Technology and Yale University.

Genesove, D., Mayer, C., 2001. Loss aversion and seller behavior: evidence from the housing market. Quarterly Journal of Economics 116, 1233–1260.

Griffin, J., Nardari, F., Stulz, R., 2007. Do investors trade more when stocks have performed well? Evidence from 46 countries. Review of Financial Studies 20, 905–951.

Grinblatt, M., Keloharju, M., 2001. What makes investors trade? Journal of Finance 56, 589–616.

Grinblatt, M., Han, B., 2005. Prospect theory, mental accounting, and momentum. Journal of Financial Economics 78, 311–339.

Grinblatt, M., Keloharju, M., 2009. Sensation-seeking, overconfidence, and trading activity. Journal of Finance 64, 549–578.

Han, B., Kumar, A., 2011. Speculative retail trading and asset prices. Unpublished working paper, University of Texas at Austin.

Heath, C., Huddart, S., Lang, M., 1999. Psychological factors and stock option exercise. Quarterly Journal of Economics 114, 601–627.

Henderson, V., 2012. Prospect theory, liquidation, and the disposition effect. Management Science, forthcoming.

Hong, H., Stein, J., 2007. Disagreement and the stock market. Journal of Economic Perspectives 12, 109–128.

Ingersoll, J., Jin, L., 2011. Intertemporal prospect theory with realization utility. Unpublished working paper. Yale University.

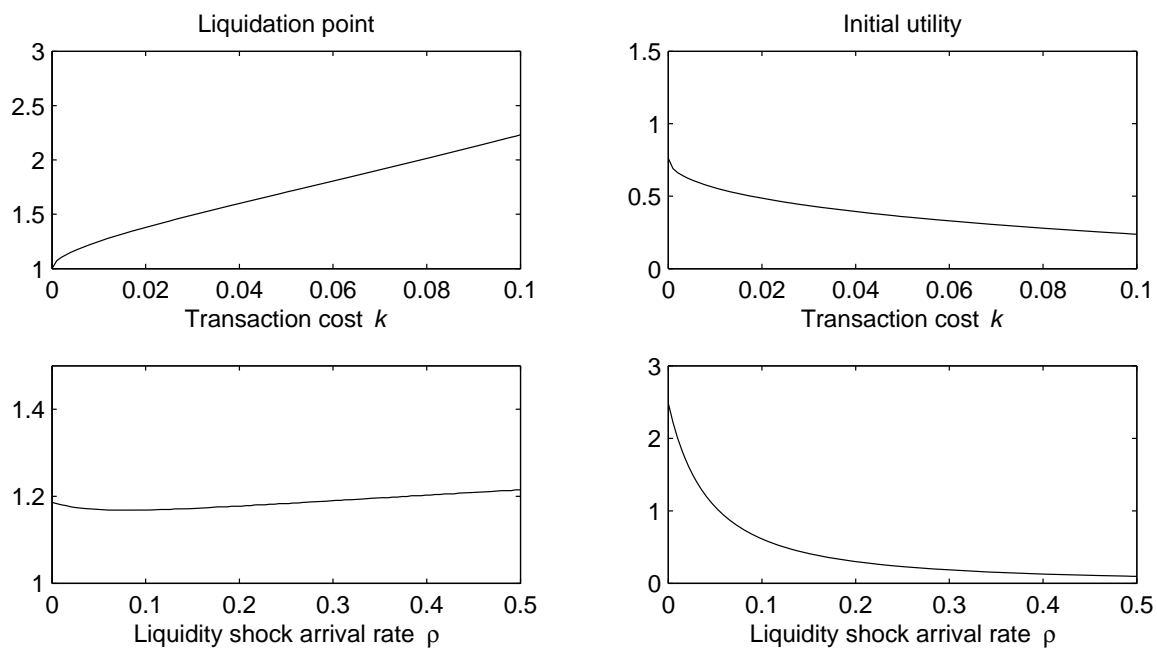Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263–292.

Kaustia, M., 2010. Prospect theory and the disposition effect. Journal of Financial and Quantitative Analysis 45, 791–812.

Kumar, A., 2009. Who gambles in the stock market? Journal of Finance 64, 1889–1933.

Markowitz, H., 1952. The utility of wealth. Journal of Political Economy 60, 151–158.

Odean, T., 1998. Are investors reluctant to realize their losses? Journal of Finance 53, 1775–1798.

Shefrin, H., Statman, M., 1985. The disposition to sell winners too early and ride losers too long. Journal of Finance 40, 777–790.

Statman, M., Thorley, S., Vorkink, K., 2006. Investor overconfidence and trading volume. Review of Financial Studies 19, 1531–1565.

Stein, J., 1995. Prices and trading volume in the housing market: a model with down-payment effects. Quarterly Journal of Economics 110, 379–406.

Thaler, R., 1999. Mental accounting matters. Journal of Behavioral Decision Making 12, 183–206.

Weber, M., Camerer, C., 1998. The disposition effect in securities trading: an experimental analysis. Journal of Economic Behavior and Organization 33, 167–184.
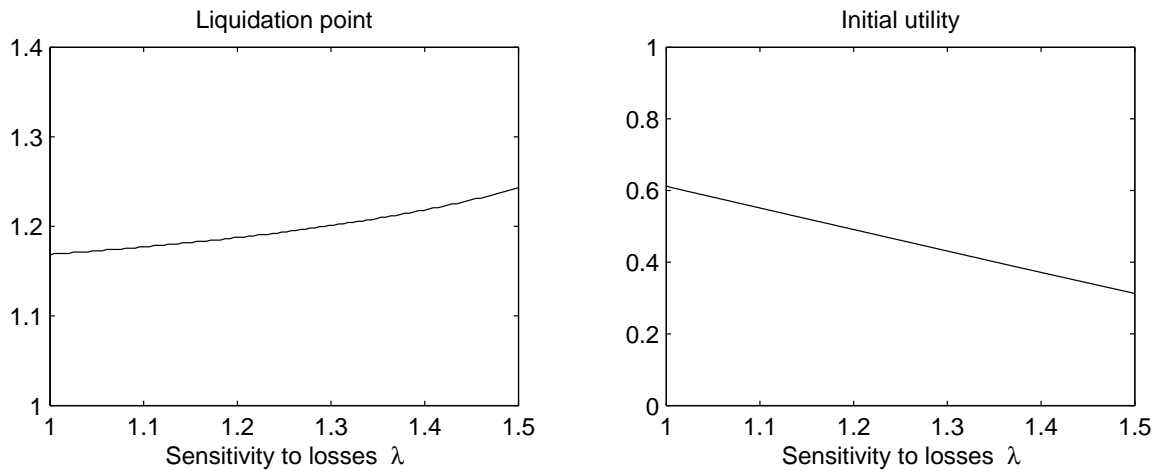
Fig. 1. Range of values of a stock's expected excess return and standard deviation for which an investor who derives utility from realized gains and losses is willing both to buy the stock and to sell it once its price reaches a sufficiently high liquidation point. The top graph corresponds to the case in which realization utility has a linear functional form. The bottom graph corresponds to the case in which realization utility has a piecewise-linear functional form, so that the investor is 1.5 times as sensitive to realized losses as to realized gains.
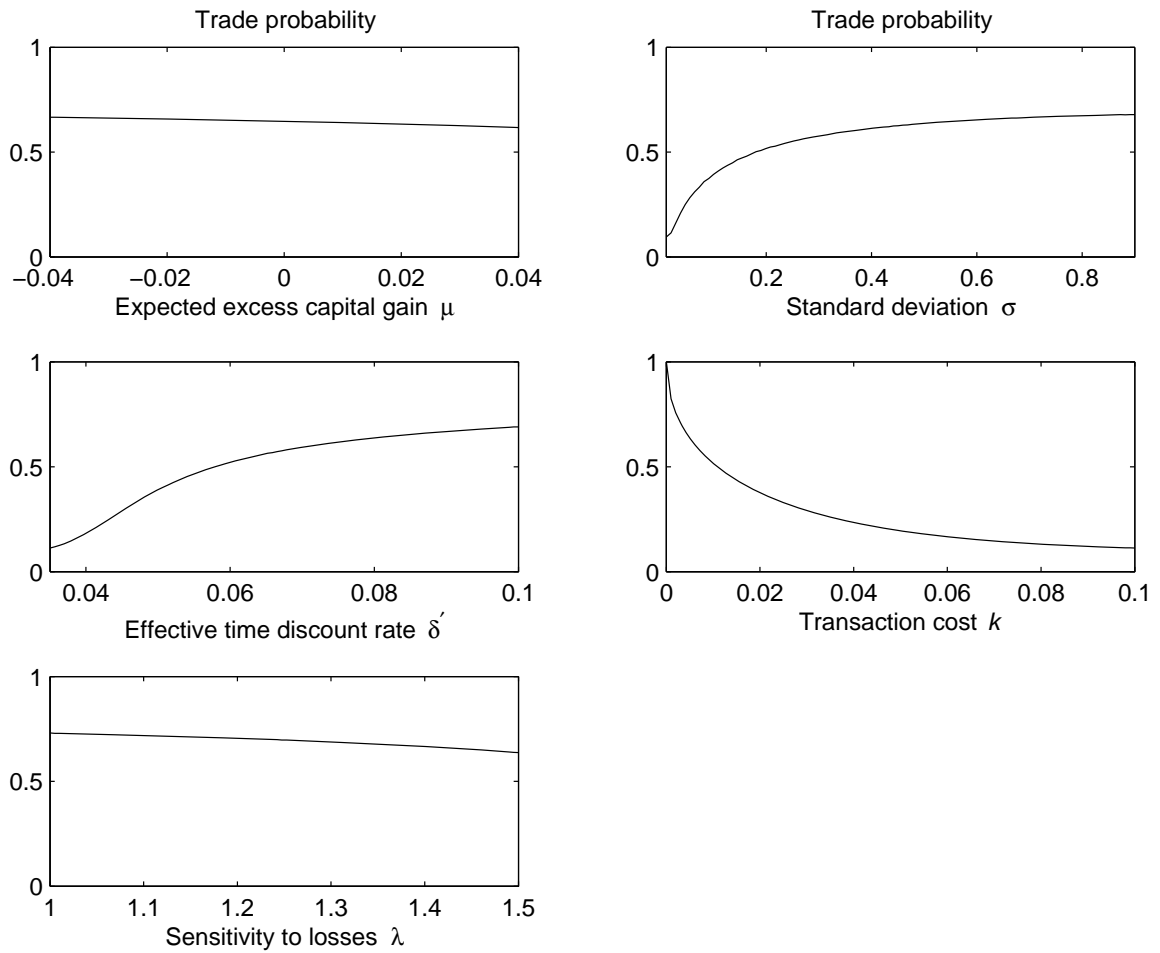
Fig. 2. Sensitivity of the liquidation point at which an investor sells a stock, and of the initial utility from buying it, to the stock's expected excess capital gain μ, its standard deviation σ, and the effective time discount rate δ'. The investor derives utility from realized gains and losses. The solid lines correspond to the case where realization utility has a linear functional form. The dashed lines correspond to the case where realization utility has a piecewise-linear functional form, so that the investor is 1.5 times as sensitive to realized losses as to realized gains.
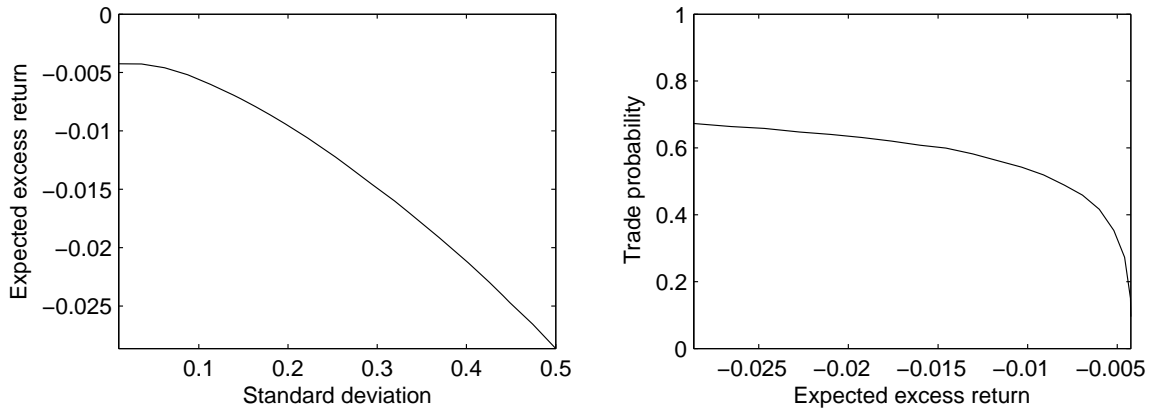
43

Fig. 3. Sensitivity of the liquidation point at which an investor sells a stock, and of the initial utility from buying it, to the transaction cost k and the arrival rate ρ of an exogeneous liquidity shock. The investor derives utility from realized gains and losses. Realization utility has a linear functional form.

Fig. 4. Sensitivity of the liquidation point at which an investor sells a stock, and of the initial utility from buying it, to λ, his relative sensitivity to realized losses as opposed to realized gains.

Fig. 5. Probability that an investor who derives utility from realized gains and losses will sell a specific stock within a year of buying it. The graphs show how this probability varies with the stock's expected excess capital gain μ, its standard deviation σ, the effective time discount rate δ', the transaction cost k, and the relative sensitivity to realized losses as opposed to realized gains λ.

46

Fig. 6. Expected return, standard deviation, and probability of sale in an economy populated by investors who derive utility from realized gains and losses. The top-left graph shows the equilibrium relationship between expected excess return and standard deviation in a cross-section of stocks. The top-right graph shows, for the same cross-section, the equilibrium relationship between a stock's expected excess return and the probability that, after buying the stock, an investor sells it within a year of purchase. Realization utility has a piecewise-linear functional form, so that investors are 1.5 times as sensitive to realized losses as to realized gains.