Realization Utility

Nicholas Barberis and Wei Xiong
Yale University and Princeton University

July 2010*

Abstract

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

*We thank Bige Kahraman for able research assistance. We are also grateful to Daniel Benjamin, Patrick Bolton, Xuedong He, Vicky Henderson, Antonio Mele, Chris Rogers, Paul Tetlock, and seminar participants at Arizona State University, Brown University, Cornell University, Harvard University, the LSE, New York University, Princeton University, the University of California at Berkeley, the University of Texas at Austin, Yale University, and the Gerzensee Summer Symposium for helpful comments. Comments are welcome at nick.barberis@yale.edu and wxiong@princeton.edu.

1 Introduction

When economists model the behavior of individual investors, they typically assume that these investors derive utility only from consumption or total wealth. In this paper, we study the possibility that investors derive utility, in part, 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 analyze a model in which he receives a burst of utility right then, at the moment of sale. The utility term depends on the size of the gain or loss realized – on the difference between sale price and 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 the way 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. This notion of investing episodes is consistent with a large body of work in psychology that shows that the brain processes ongoing experiences by segmenting them into episodes (Zacks and Swallow, 2007).

The second cognitive process that, in our view, underlies realization utility has to do with the way 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 wealth 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 wealth than he had at the start.

In summary, then, 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 circumstances. 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. Indeed, it is probably most relevant for the least sophisticated individual investors. We also expect realization utility to play a larger role when an asset's purchase price is more salient: it is easier to declare that an investment is a success when success is easier to measure. Realization utility 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 from realizing a gain or loss 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 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 develop a model of realization utility, one that is sophisticated enough to capture many features of actual trading, but also simple enough to allow for an analytical solution. We then link the model to a wide range of applications. Finally, we lay out some new predictions. 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 framework 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 he sells a position in stock, he receives a jolt 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 realization 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, in our model, an investor who is holding a position in stock will voluntarily sell this position only if the stock price rises sufficiently far above the purchase price. We look at how the "liquidation point" – the percentage gain in price, relative to purchase price, 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 also allows us to compute the probability that the investor sells the stock within a certain amount of time after the initial purchase. We look at how this probability depends on the aforementioned factors.

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-*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 by a liquidity shock. Any realized loss therefore lies in the distant, heavily discounted future and does not scare the investor very much today. Overall, then, the investor prefers more volatility to less.

We use our model to link realization utility to a wide range of financial phenomena. Among the applications we discuss are the subpar trading performance of individual investors (Barber and Odean, 2000; Barber et al., 2009), the disposition effect (Odean, 1998), the greater turnover in bull markets than in bear markets (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 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 puzzling tendency of individual investors to sell stocks that have risen in value, rather than fallen in value, since purchase. Our model shows that realization utility, in combination with a positive time discount rate, predicts 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 "relabelling" of the disposition effect. To 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 possible 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 be affected by realization utility – will have a much greater propensity to sell a stock once it 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 very little selling as the stock approaches \$45 for the second time – any realization utility investors with liquidation points of \$45 or lower *will have sold already* when the stock first approached \$45 – but as soon as the stock moves above the historical high of \$45, realization utility investors with liquidation points higher than \$45 will suddenly start selling. In line with the recent 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 we should study models in which investors derive utility from *paper* gains and losses. In Benartzi and Thaler (1995), for example, investors derive utility from fluctuations in the value of their financial wealth, while in Barberis, Huang, and Santos (2001) and Barberis and Huang (2001), 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 less attention, probably because it is seen as a more radical departure from the standard model than simply defining utility over paper gains and losses. The concept first appears in an important paper by Shefrin and Statman (1985). Among several other contributions, these authors point out, with the help of a simple 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 Shefrin and Statman's (1985) 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.

Other authors also discuss realization utility. In an article on mental accounting, 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)." And Odean (1999) remarks that "if the investor is psychologically motivated, he may wish to avoid realizing losses."

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 make just one observation: that, consistent with Shefrin and Statman (1985), the investor often exhibits a disposition effect. They do not study any other implications of realization utility, nor do they link it to any other applications.[1]

[1]Barberis and Xiong (2009) do not say very much about realization utility because it is not their main

In this paper, we present 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 portfolio choice and a model of asset pricing. We document, for the first time, several basic implications of realization utility. And we discuss eight potential applications, rather than just one. Finally, we list several new predictions.

In Section 2, we present the baseline model, one that assumes a linear functional form for realization utility. In Section 3, we use a piecewise-linear functional form. In Section 4, we show how realization utility can be embedded in a full equilibrium. Section 5 discusses a range of applications and new 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 sells a specific asset. It is therefore natural that we 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, however, 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 return; and N risky assets indexed by $i \in \{1, \ldots, N\}$. To avoid excessive notation, we use the risk-free asset as the numeraire for all asset prices and for the investor's wealth. Specifically, the price of risky asset i, $S_{i,t}$, denominated in the risk-free asset, follows

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

where $Z_{i,t}$ is a Brownian motion and where, for $i \neq j$, $dZ_{i,t}$ and $dZ_{j,t}$ may be correlated.

focus. Rather, their paper is primarily about the trading behavior of an investor who derives prospect theory utility from *paper* gains and losses.

Risky asset i also pays a stream of dividends $D_{i,t}$, where

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

The asset's instantaneous expected excess return – throughout the paper, "excess" means over and above the risk-free rate – is therefore $\alpha + \mu$: in words, 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 across all the risky assets. The most natural application of our model is to understanding how individual investors trade stocks in their brokerage accounts. We therefore often refer to the risky assets as stocks.

The dividends $D_{i,t}$ do not play a significant role in the partial equilibrium analysis of 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 our model in a full equilibrium framework. To prevent the dividends from unnecessarily complicating our 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.[2]

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. We also suppose that, if the investor sells his position in a stock at time t, he pays a proportional transaction cost, kW_t, $0 \leq k < 1$, where W_t is time t wealth. The investor's wealth therefore evolves according to

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

where $\theta_{i,t}$ takes the value 1 if he is holding stock i at time t, and 0 otherwise; and where l_t takes the value 1 if he sells a stock at time t, and 0 otherwise. Note that if $\theta_{i,t} = 1$ for some i and t, then $\theta_{j,t} = 0$ for all $j \neq i$.[3]

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

[2] We have not been able to find a tractable way of incorporating a *concave* consumption utility term into our model. However, we *have* been able to study a preference specification that combines realization utility with concave consumption utility in a *two-period* version of our model. In this two-period model, we find that, so long as the weight on consumption utility is not too large, a preference specification that combines realization utility with concave consumption utility delivers similar predictions to a specification that combines realization utility with linear consumption utility.

[3] The assumption that, at each time, the investor allocates all of his wealth either to the risk-free asset or to one of the N stocks is not as strong as it initially appears to be. We will see later that the behavior assumed here is *optimal* under a much weaker assumption that we will specify explicitly.

words, the amount of money the investor put into the time t stock position at the time he bought it. Formally, if $\theta_{i,t} = 1$,

$$B_t = W_s, \text{ where } s = \min\{\tau \in [0, t] : \theta_{i,\tau'} = 1 \text{ for all } \tau' \in [\tau, t]\}. \quad (5)$$

As with the investor's wealth, the cost basis B_t is measured using the risk-free asset as numeraire.

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

$$u((1 - k)W_t - B_t). \quad (6)$$

The argument of the utility term is the size of 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 a linear functional form,

$$u(x) = x. \quad (7)$$

We emphasize that the investor only receives the burst of utility in (6) if he switches 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. Realization utility is associated with the *completion* of a transaction. It is hard to argue that the sale of a stock represents a completed transaction if, after selling the stock, the investor immediately buys it back.[4]

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 and exits the asset markets. We think of this shock as capturing an unexpected 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. The liquidity shock 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.

Suppose that, at time t, the investor's wealth is allocated to a stock. The investor's value function depends on two things: on the current asset value, W_t, and on the asset's cost basis, B_t. We therefore denote it as $V(W_t, B_t)$. Since the utility functions in (3) and (7)

[4] We assume that the investor does not incur a transaction cost when he sells the risk-free asset. Given our definition of the cost basis, the realized gain or loss from selling the risk-free asset is therefore always zero. As a result, the investor only receives realization utility when he sells stock, not when he sells the risk-free asset.

are homogeneous of degree one, and since the prices of the risky assets all follow a geometric Brownian motion, 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 from 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 condition (9) holds for some positive W, then it holds for all positive W. Later, we will compute the range of parameter values for which condition (9) holds. For now, we note that, given a positive time discount rate, condition (9) implies two things. First, it implies that, at time 0, the investor allocates his wealth to a stock: since the risk-free asset generates no utility flows, he allocates to a stock as early as possible. Second, and using a similar logic, condition (9) implies that, if, at any time $t > 0$, the investor sells a position in 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

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

subject to equations (3), (4), (5), and (7). $I_{\{\}}$ is an indicator function that takes the value 1 if the condition in the curly brackets is met, and 0 otherwise. The parameter δ is the time discount rate. To ensure that the investor does not hold his time 0 stock position forever, without selling it, we impose the following parameter restriction:

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

In simple terms, the restriction ensures that the expected stock return is not too high.

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 summarizes the solution to the decision problem in (10). It states that, if the investor buys a stock, his optimal strategy is to sell it 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 percentage change in value, between the time of purchase and time t, of the risky asset the investor is holding at time t – plays an important role in the solution.

Proposition 1: Unless forced to exit the stock market by a liquidity shock, an investor with the decision problem in (10) will sell a position in stock once 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 investor's value function is $V(W_t, B_t) = B_t U(g_t)$, where

$$U(g_t) = \begin{cases} a g_t^{\gamma_1} + \frac{\alpha\beta + \rho(1-k)}{\rho+\delta-\mu} g_t - \frac{\rho}{\rho+\delta} & \text{if} \quad g_t < g_* \\ \\ (1-k)(1+U(1))g_t - 1 & \text{if} \quad g_t \geq g_* \end{cases}, \tag{13}$$

where

$$\gamma_1 = \frac{1}{\sigma^2}\left[\sqrt{\left(\mu - \frac{1}{2}\sigma^2\right)^2 + 2(\rho+\delta)\sigma^2} - \left(\mu - \frac{1}{2}\sigma^2\right)\right] > 0 \tag{14}$$

and

$$a = \frac{\delta}{g_*^{\gamma_1}(\gamma_1 - 1)(\rho+\delta)}. \tag{15}$$

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

$$(\gamma_1 - 1)\left(1 - \frac{k(\rho+\delta)(\rho + \frac{\alpha\beta}{1-k})}{\delta(\rho+\delta-\mu)}\right) g_*^{\gamma_1} - \frac{\gamma_1}{1-k} g_*^{\gamma_1 - 1} + 1 = 0. \tag{16}$$

We prove the proposition in the Appendix. The proof proceeds by conjecturing that the investor sells his stock position once g_t exceeds some value g_*; by constructing the value function, first for the region below g_* and then for the region above g_*; by requiring that the value function is continuous and continuously differentiable at g_*; and finally, by verifying that the constructed value function is indeed optimal.

In summary, then, the investor's optimal strategy takes one of two forms, each of which has a simple structure. 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, the investor buys a stock at time 0 and voluntarily sells it only if it reaches a sufficiently high liquidation point, at which point he immediately invests the proceeds in another stock, and so on. 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.[5]

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 at the same time. Suppose that he starts at time 0 with wealth of mW_0 and spreads this wealth across m stocks, investing W_0 in each one. 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 stock 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.

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 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 value function for the first strategy, $W_0U(1)$, is the same as the value function for the second strategy, namely $W_0U(1)/2 + W_0U(1)/2$.[6]

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 values of the model parameters 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. When assigning parameter values, we have in mind our model's most natural application, namely stock trading in brokerage accounts by individual

[5] Since $g_* \geq 1$, the term $U(1)$ which appears in the second row of equation (13) can be directly obtained from the first row of equation (13). Specifically, it equals $a + (\alpha\beta + \rho(1-k))/(\rho + \delta - \mu) - \rho/(\rho + \delta)$.

[6] Suppose that we make the weak assumption that, if the investor is indifferent to diversification, he simply invests in a single stock. Since the investor *is* indifferent to diversification at time 0, he therefore invests in a single stock at that time. The linearity of $V(W, B)$ in the cost basis B means that, even if he is allowed to engage in a partial sale of the stock, he will not want to. Instead, he will either sell his entire stock position or else keep it intact. If he sells, he is back in the same situation as at time 0: since he is indifferent to diversification, he invests the proceeds of the sale in a single stock. Our earlier assumption that the investor, at all times, allocates his wealth to either the risk-free asset or to one of the N stocks, is therefore not as strong as it initially appears to be. Rather, it follows directly from the linear structure of our model and from one much weaker assumption: that, if the investor is indifferent to diversification, he invests in a single stock.

investors.

The shaded area in the top graph in Figure 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 the investor is willing to sell the stock at a finite liquidation point.

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 $k = 0.005$, which is of a similar order of magnitude to the transaction cost estimated by Barber and Odean (2000) for discount brokerage customers. We set $\rho = 0.1$, so that the probability of a liquidity 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. Given these parameter values, we choose a 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.[7]

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 simple. So long as σ is sufficiently high, even a negative expected return stock has a non-negligible chance of reaching the liquidation point g_*, at which time the investor can enjoy realizing the gain. Of course, more likely than not, the stock will lose value. However, since the investor does 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, heavily discounted future and does not scare the investor very much today. Overall, then, investing in stock, even if it has a negative expected excess return, can be better than investing in the risk-free asset, which offers zero utility.[8]

Figures 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,

[7]While the discount rate in our model is higher than in many consumption-based pricing models, the approach we take in setting the discount rate is the *same* as in those other models. Specifically, researchers working with consumption-based models typically choose the discount rate so as to match a feature of the data they are interested in – in most cases, this is the real interest rate. In the same way, we choose the discount rate to match a feature of the data that we are interested in, namely the frequency of trading in brokerage accounts.

[8]Note that, when $\rho = 0$, in other words, when there is no liquidity shock, Proposition 1 tells us that the investor is willing to buy a stock at time 0 for *any* values of the expected excess return $\alpha + \mu$ and standard deviation σ. The reason is that, when $\rho = 0$, the investor only sells a stock if it reaches its liquidation point, which, for $k > 0$, exceeds 1; in particular, he never sells a stock at a loss, for any reason. Since owning a stock never leads to any bursts of negative utility, the investor is always willing to buy one at time 0, whatever the values of $\alpha + \mu$ and σ.

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 a 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 Figures 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 Figure 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 the form of realization utility is linear, the investor is risk-*seeking*. While this is initially surprising, there is a simple intuition for it. A highly volatile stock offers the chance of a significant gain, which the investor can enjoy realizing. Of course, it also offers the chance of a significant loss. 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, heavily discounted future and does not scare the investor very much today. Overall, then, the investor prefers more volatility to less. In mathematical terms, this prediction is a consequence of the fact that, while instantaneous utility is linear, the value function $U(g_t)$ in (13) is convex.[9] A similar intuition explains why, in the middle-left graph, the liquidation point is increasing in volatility.

The trading patterns we have just described – the buying of low expected return stocks and the preference for volatility – 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, instead of thinking about their investment performance in the correct way, using overall portfolio return, think about it as a series of investing episodes and, as a result, allow realization utility to affect their trading. What Figures 1 and 2 demonstrate is that an investor who truly believes that a decline in the value of a stock "is only a loss if I realize it" can be drawn into stocks with low expected returns and high volatility. We discuss some evidence consistent with this prediction in Section 5.

[9]The parameter restriction in (11) implies $\gamma_1 > 1$ and $a > 0$, which, in turn, implies the convexity of $U(\cdot)$.

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

The bottom-right graph shows that initial utility is relatively insensitive to the discount rate. There are two opposing forces at work here. On the one hand, a lower value of δ means that future utility flows are discounted at a lower rate, thereby raising initial utility. On the other hand, since an investor with a low discount rate sets a high liquidation point, he may have to endure the unpleasant scenario whereby a stock initially rises quite high, although not as high as the liquidation point, and then falls, generating a paper loss from which he is eventually forced to exit by a liquidity shock. By contrast, an investor with a high discount rate and hence a low liquidation point is less likely to experience such a scenario. These opposing effects lead to a relatively flat relationship between initial utility and the discount rate.

The top graphs in Figure 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 Figure 3 suggests that, in this case, the liquidation point is $g_* = 1$. It is straightforward to check that, when $k = 0$, equation (16) 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 Figure 3 show how the liquidation point and initial utility depend on ρ, the intensity of the liquidity shock. The bottom-left graph shows that 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 is likely to 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 from which he might be forced to exit 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 the investor will be forced to exit the stock market with a painful loss.[10]

Several of the implications of realization utility that we have described – the investor's willingness to buy stocks with a low expected return, the fact that he only voluntarily sells stocks at a gain, and the fact that he is risk-seeking – can also be seen 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, surprisingly, simpler than in a two-period model. Indeed, we saw earlier that the investor's optimal strategy has a *very* simple structure: 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.

Why is the structure of the solution simpler in an infinite horizon model? It is due to the fact that, in such a model, the environment is *stationary*. For example, the investor's value function does not depend explicitly on time, t, but only on the state variable g_t. In a two-period model, however, the environment is non-stationary and so the investor's trading strategy, while similar to that in our model, has a somewhat more complex structure.

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:

$$u(x) = \begin{cases} x & \text{if } x \geq 0 \\ \lambda x & \text{if } x < 0 \end{cases}, \qquad \lambda > 1, \tag{17}$$

where λ controls the relative sensitivity to realized losses as opposed to realized gains.[11]

[10] What do the solid lines in Figure 2 and in the top panel of Figure 3 look like when there is no liquidity shock? We find that, if we recompute these lines using the same benchmark parameter values for α, μ, σ, k, δ, and β, but with a ρ of 0 rather than 0.1, there is no significant change. The relationships between the liquidation point and initial utility on the one hand, and the parameters μ, σ, δ, and k on the other, remain qualitatively the same and quantitatively quite similar.

[11] 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 involves different psychological factors. 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.

The investor's decision problem is now

$$
\begin{aligned}
V(W_t, B_t) &= \max_{\tau \geq t} \; E_t \{ \int_t^{\min\{\tau,\tau'\}} e^{-\delta(s-t)} v(D_{i,s}) ds \\
&\quad + e^{-\delta(\tau-t)} [u((1-k)W_\tau - B_\tau) + V((1-k)W_\tau, (1-k)W_\tau)] I_{\{\tau<\tau'\}} \\
&\quad + e^{-\delta(\tau'-t)} u((1-k)W_{\tau'} - B_{\tau'}) I_{\{\tau \geq \tau'\}} \}, \qquad (18)
\end{aligned}
$$

subject to (3), (4), (5), and (17). 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 (17).

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 (18) will sell a position in stock once 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 investor's value function is $V(W_t, B_t) = B_t U(g_t)$, where

$$
U(g_t) = \begin{cases}
b g_t^{\gamma_1} + \frac{\alpha\beta+\rho\lambda(1-k)}{\rho+\delta-\mu} g_t - \frac{\rho\lambda}{\rho+\delta} & \text{if} \quad g_t \in \left(0, \frac{1}{1-k}\right) \\
c_1 g_t^{\gamma_1} + c_2 g_t^{\gamma_2} + \frac{\alpha\beta+\rho(1-k)}{\rho+\delta-\mu} g_t - \frac{\rho}{\rho+\delta} & \text{if} \quad g_t \in \left(\frac{1}{1-k}, g_*\right) \\
(1-k) g_t (1 + U(1)) - 1 & \text{if} \quad g_t \in (g_*, \infty)
\end{cases}, \qquad (19)
$$

where γ_1 is defined in equation (14), where

$$
\gamma_2 = -\frac{1}{\sigma^2}\left[\sqrt{\left(\mu - \frac{1}{2}\sigma^2\right)^2 + 2(\rho+\delta)\sigma^2} + \left(\mu - \frac{1}{2}\sigma^2\right)\right] < 0, \qquad (20)
$$

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

$$
c_2 = \frac{(\lambda-1)\rho(1-k)^{\gamma_2}(\mu\gamma_1 - \rho - \delta)}{(\gamma_1-\gamma_2)(\rho+\delta-\mu)(\rho+\delta)} \qquad (21)
$$

$$
(\gamma_1 - 1) c_1 g_*^{\gamma_1} + (\gamma_2 - 1) c_2 g_*^{\gamma_2} = \frac{\delta}{\rho+\delta} \qquad (22)
$$

$$
c_1 \left(\frac{1}{1-k}\right)^{\gamma_1} + c_2 \left(\frac{1}{1-k}\right)^{\gamma_2} = b\left(\frac{1}{1-k}\right)^{\gamma_1} + \frac{(\lambda-1)\mu\rho}{(\rho+\delta-\mu)(\rho+\delta)} \qquad (23)
$$

$$
c_1 g_*^{\gamma_1} + c_2 g_*^{\gamma_2} + \frac{k\alpha\beta + (1-k)(\mu-\delta)}{\rho+\delta-\mu} g_* + \frac{\delta}{\rho+\delta} = (1-k) g_* \left(b + \frac{\rho\lambda(\mu - k\rho - k\delta)}{(\rho+\delta)(\rho+\delta-\mu)}\right) \qquad (24)
$$

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 equation (21) to find c_2; we then obtain c_1 from equation (22); we then use equation (23) to find b; finally, equation (24) allows us to solve for the liquidation point g_*.

Results

The shaded area in the lower graph in Figure 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 both willing to buy a stock at time 0, so that condition (9) holds, and also to sell it 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: 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 a negative expected return 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 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.

The graphs in Figure 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

$$\begin{aligned} (\alpha, \mu, \sigma, k) &= (0.015, 0.015, 0.5, 0.005) \\ (\delta, \rho, \beta) &= (0.08, 0.1, 1). \end{aligned} \tag{25}$$

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 from which he might be forced to exit 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 Figure 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,

$$\begin{aligned} (\alpha, \mu, \sigma, k) &= (0.015, 0.015, 0.5, 0.005) \\ (\delta, \rho, \lambda, \beta) &= (0.08, 0.1, 1.5, 1). \end{aligned} \tag{26}$$

Recall how the calculations for the solid lines in Figure 2 differ from those for the dashed lines: the solid lines correspond to linear realization utility, so that $\lambda = 1$; the dashed lines assume $\lambda = 1.5$. The dashed lines show 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 Figure 2 deserves particular attention. It shows that, for the benchmark values in (26), initial utility $U(1)$ is still increasing in stock volatility σ. Put differently, even though the functional form for realization utility is now concave, the investor is still risk-seeking. If the sensitivity to losses λ or the liquidity shock intensity ρ rise significantly, however, this relationship reverses: initial utility becomes a decreasing function of σ and the investor is risk averse, not risk-seeking.

4 Asset Pricing

In Sections 2 and 3, we studied realization utility in a partial equilibrium setting. In this section, we show how it can be embedded in a full equilibrium. This will help us understand its implications for asset prices. Of course, if realization utility is to affect prices, *many* investors must care about it. It is hard to know, based on introspection alone, whether this is the case. Perhaps the best way to find out is to derive the pricing implications of realization utility and to see if this sheds light on puzzling facts and leads to new predictions that can be tested and confirmed.

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 net return of zero. The risky stocks are in limited supply. The dividend process for stock i is

$$\frac{dD_{i,t}}{D_{i,t}} = \mu_i dt + \sigma_i dZ_{i,t}, \tag{27}$$

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 determined in equilibrium. We hypothesize that

$$S_{i,t} = \frac{1}{\alpha_i} D_{i,t}, \tag{28}$$

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

$$\frac{dS_{i,t}}{S_{i,t}} = \mu_i dt + \sigma_i dZ_{i,t}. \tag{29}$$

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

In this economy, the equilibrium conditions are

$$V_i\left(W, W\right) = 0, \qquad i = 1, \ldots, N, \tag{30}$$

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 equations (30) the appropriate equilibrium conditions? Note that, under the conditions in (30), 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 (30) mean that they are happy to then be assigned to the risk-free asset. Finally, the conditions in (30) 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 stock i.[12]

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

$$\begin{aligned} V_i\left(W_t, B_t\right) &= \max_{\tau \geq t} \; E_t\Big\{\int_t^{\min\{\tau,\tau'\}} e^{-\delta(s-t)} v(D_{i,s}) ds \\ &\quad + e^{-\delta(\tau-t)} u((1-k_i)W_\tau - B_\tau) I_{\{\tau<\tau'\}} + e^{-\delta(\tau'-t)} u((1-k_i)W_{\tau'} - B_{\tau'}) I_{\{\tau \geq \tau'\}}\Big\}, \end{aligned} \tag{31}$$

subject to equations (3), (5), (17), and

$$\frac{dW_s}{W_s} = \mu_i ds + \sigma_i dZ_{i,s}, \qquad t \leq s < \min\{\tau, \tau'\}, \tag{32}$$

where τ' is the random future time at which a liquidity shock arrives. This differs from the decision problem in (18) in that it imposes the market clearing condition (30): after selling his stock holdings at time τ, the investor's future value function is zero. We summarize the solution to the decision problem in (31) in the following proposition. The proof is in the Appendix.

[12]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.

Proposition 3: Unless forced to exit the stock market by a liquidity shock, an investor with the decision problem in (31) will sell a position in stock once 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 investor's value function when holding stock i at time t is $V_i(W_t, B_t) = B_t U_i(g_t)$, where

$$U_i(g_t) = \begin{cases} b g_t^{\gamma_1} + \frac{\alpha_i\beta+\rho\lambda(1-k_i)}{\rho+\delta-\mu_i} g_t - \frac{\rho\lambda}{\rho+\delta} & \text{if} \quad g_t \in (0, \frac{1}{1-k_i}) \\ c_1 g_t^{\gamma_1} + c_2 g_t^{\gamma_2} + \frac{\alpha_i\beta+\rho(1-k_i)}{\rho+\delta-\mu_i} g_t - \frac{\rho}{\rho+\delta} & \text{if} \quad g_t \in (\frac{1}{1-k_i}, g_*) \\ (1-k_i)g_t - 1 & \text{if} \quad g_t \in (g_*, \infty) \end{cases}, \tag{33}$$

where γ_1 and γ_2 are given by

$$\gamma_1 = \frac{1}{\sigma_i^2}\left[\sqrt{\left(\mu_i - \frac{1}{2}\sigma_i^2\right)^2 + 2(\rho+\delta)\sigma_i^2} - \left(\mu_i - \frac{1}{2}\sigma_i^2\right)\right] > 0 \tag{34}$$

$$\gamma_2 = -\frac{1}{\sigma_i^2}\left[\sqrt{\left(\mu_i - \frac{1}{2}\sigma_i^2\right)^2 + 2(\rho+\delta)\sigma_i^2} + \left(\mu_i - \frac{1}{2}\sigma_i^2\right)\right] < 0, \tag{35}$$

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

$$c_2 = \frac{(\lambda-1)\rho(1-k_i)^{\gamma_2}(\mu_i\gamma_1 - \rho - \delta)}{(\gamma_1-\gamma_2)(\rho+\delta-\mu_i)(\rho+\delta)} \tag{36}$$

$$(\gamma_1 - 1)c_1 g_*^{\gamma_1} + (\gamma_2 - 1)c_2 g_*^{\gamma_2} = \frac{\delta}{\rho+\delta} \tag{37}$$

$$c_1\left(\frac{1}{1-k_i}\right)^{\gamma_1} + c_2\left(\frac{1}{1-k_i}\right)^{\gamma_2} = b\left(\frac{1}{1-k_i}\right)^{\gamma_1} + \frac{(\lambda-1)\mu_i\rho}{(\rho+\delta-\mu_i)(\rho+\delta)} \tag{38}$$

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

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 exogeneously given. To determine α_i, we require that the value function satisfy the conditions in (30), namely $V_i(W, W) = 0$, or equivalently, $U_i(1) = 0$. The parameter α_i therefore satisfies

$$b + \frac{\alpha_i\beta + \rho\lambda(1-k_i)}{\rho+\delta-\mu_i} - \frac{\rho\lambda}{\rho+\delta} = 0. \tag{40}$$

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

In Section 5.2, we use the model described in this section to illustrate some implications of realization utility for asset prices. We emphasize that conditions (30) 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. A useful direction for future research would be to study asset prices in an economy with both realization utility investors *and* expected utility investors. We conjecture that, in such an economy, the expected utility investors would 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 who care more about realization utility. In Section 5.3, we use this idea as the basis for some new predictions.

5 Applications

Our model may be helpful for thinking about a number of financial phenomena. We now discuss some of these potential applications. We divide the applications into those that relate to investor trading behavior (Section 5.1) and those that relate to asset prices (Section 5.2). In Section 5.3, we discuss some of our model's testable predictions.

5.1 Investor trading behavior

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 (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 subsequently 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 they hold on to. Odean (1998) also considers other potential explanations based on taxes, rebalancing, and transaction costs, but argues that all of them fall short.

Our analysis shows that a model that combines realization utility with a positive time discount rate predicts a strong disposition effect. Unless forced to sell by a liquidity shock, the investor in our model *only* sells stocks trading at a gain, never a stock trading at a loss.

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 relabelling of the disposition effect. To 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 (1995). 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 (1995) 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.

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 a simple explanation for this post-transaction-cost underperformance. Under this view, the investors in Barber and Odean's (2000) sample experience realization utility. While they are aware that they underperform the benchmarks on average, they are compensated for this underperformance by the occasional bursts of positive utility they receive when they realize gains.

It is straightforward to compute the probability that the investor in our model sells a stock within any given interval after the time of 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, $g_0 = 1$. When g_t reaches an upper barrier $g_* > 1$ or when a liquidity shock arrives, he sells the stock. To 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 passes g_* in $(0, s)$ or that a liquidity shock arrives 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:

$$\begin{aligned} G(s) &= 1 - e^{-\rho s} \\ &\quad + e^{-\rho s}\left[N\left(\frac{-\ln g_* + \left(\mu - \frac{\sigma^2}{2}\right)s}{\sigma\sqrt{s}}\right) + e^{\left(\frac{2\mu}{\sigma^2}-1\right)\ln g_*} N\left(\frac{-\ln g_* - \left(\mu - \frac{\sigma^2}{2}\right)s}{\sigma\sqrt{s}}\right)\right]. \end{aligned} \tag{41}$$

The expression in the square parentheses in (41) is the probability that g_t reaches g_* in the interval $(0, s)$. With this information in hand, it is easy to interpret equation (41). 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_*.

Figure 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 equations (21)-(24) 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, while keeping the remaining parameters fixed at their benchmark values

$$\begin{aligned}(\alpha, \mu, \sigma, k) &= (0.015, 0.015, 0.5, 0.005) \\ (\delta, \rho, \lambda, \beta) &= (0.08, 0.1, 1.5, 1). \qquad (42)\end{aligned}$$

Some of the results in Figure 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 graphs with μ and σ on the horizontal axis 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. However, as we saw in Figure 2, the liquidation point g_* itself goes up as μ and σ go up, thereby *lowering* the chance that g_* will be reached within the year-long interval. Without computing $G(1)$ explicitly, we cannot tell which factor will dominate.

The top graphs in Figure 5 show that, interestingly, a different factor dominates in each of the two cases. 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 graph with λ on the horizontal axis shows that trading frequency declines as the investor's sensitivity to losses rises. The intuition is that, 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 from which he might be forced, by a liquidity shock, to make a painful exit.

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. Figure

5 shows that, for the benchmark parameter values, our model predicts a trading frequency that is of a similar order of magnitude. When $\sigma = 50\%$, 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 volume of trading in financial markets. Rather, we mean something narrower: that it can help us understand why individuals 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 care about realization utility would lead to an increase in overall trading volume.

Underperformance *before* transaction costs

Some studies find that the average individual investor underperforms benchmarks even before transaction costs (Odean, 1999; Barber et al., 2009). Our model may be able to shed light on this. The key insight is that, as discussed in Section 2, an investor who cares about realization utility is 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 care about realization utility; and institutions, who do not. Since individuals care about realization utility, they are more willing than are 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 market benchmarks, again before transaction costs. This prediction is broadly consistent with the available evidence.

Trading in rising and falling markets

Researchers have found that, across many different asset classes, there is more trading in rising markets than in falling markets (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 who move out of the risk-free asset and into the stock market. 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.

The effect of historical highs on the propensity to sell

Our model predicts that there will be more trading activity in rising markets, but it can also make more precise predictions about the dynamics of trading. For example, it predicts that individual investors – the investor group that is more likely to care about realization utility – will have a much higher propensity to sell a stock once the stock price moves above its *historical high*.

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 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 individuals. 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 during 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. As the stock moves above $45 on March 31st, however, investors targeting liquidation points higher than $45 will suddenly 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 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.[13]

Our model is also consistent with another of Grinblatt and Keloharju's (2001) findings: that the historical high over the past *six months* or *year* is a less significant predictor of the propensity to sell. To see this, consider, as before, a stock that rises from $30 on January 1st to $45 by February 28th. Suppose that the stock then declines and trades below $45 for *several months*, and that only in November does it finally start rising again, passing through the previous high of $45 and continuing upwards.

In this case, we would expect to see a much milder increase in selling intensity in November as the stock passes through its previous high. There are now many months between the initial high of $45 in February and the subsequent $45 mark in November. It is likely that, during this long stretch, many *new* investors will have bought the stock. If some of these investors target liquidation points lower than $45, they will sell in November as the stock rises towards $45. There will therefore be sustained selling in November not only after the stock passes through $45, but also before it gets there. As a result, there will be only a mild shift in selling intensity, if any at all, as the stock passes through its historical high in November.[14]

5.2 Asset pricing

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

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: highly volatile stocks subsequently earn low average returns. This is true not only for the U.S. stock market but for most international stock markets as well.

[13]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 we can obtain Grinblatt and Keloharju's (2001) result from 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.

[14]The equilibrium model of Section 4 allows us to make precise the statement that "during this long stretch, many *new* investors will have bought the stock." During the long stretch from February to November, liquidity shocks will force many investors to sell their positions in the stock. Those positions are bought by other realization utility investors who move from the risk-free asset into the stock market. It is therefore indeed the case that "many new investors will have bought the stock."

The finding we have just described 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. Nor can the result be fully explained using a model that combines differences of opinion with short-sale constraints: the effect persists even after controlling for differences of opinion using dispersion in analyst forecasts.

Our model offers a novel explanation. The key insight comes from the middle-right graph in Figure 2: the finding that, holding other parameters constant, initial utility is increasing in a stock's volatility. This result suggests that highly volatile stocks may experience heavy buying pressure from investors who care about realization utility. These stocks may therefore become overpriced and, as a result, may earn a low average return.

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

$$(\delta, \rho, \lambda, \beta) = (0.08, 0.1, 1.5, 1), \tag{43}$$

and assume that the excess dividend growth rate and 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 the equilibrium condition in (40) 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 markets to clear.[15]

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

The top-left graph also shows that, for the parameter values in (43), stocks earn *negative* average excess returns. We emphasize that this is not necessarily inconsistent with the positive historical aggregate equity premium. This is because our analysis only applies to stocks that are primarily held by investors who care about realization utility, which, under the assumption that individuals care more about realization utility than institutions do, means stocks that are primarily held by individuals. 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

[15] Since we are using the risk-free asset as the numeraire throughout the paper, μ is the *excess* dividend growth rate. A negative value of μ does not therefore necessarily mean that the dividend growth rate is negative – just that it is below the risk-free rate. Since all 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 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 on risky assets. The expected excess return is $\alpha + \mu$. This can be positive even if μ is negative.

stocks held by individual investors. This is exactly the finding of Han and Kumar (2008), a paper that we discuss further in Section 5.3.[16]

The high turnover 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 internet 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. Moreover, it predicts that this phenomenon should occur for assets whose value is especially uncertain.

Suppose that the uncertainty about an asset's value goes up, thereby pushing up σ, the standard deviation of returns. As noted earlier, investors who care about 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 Figure 5 shows that, as σ goes up, the probability that an investor will trade the asset *also* goes up: simply put, a more volatile stock reaches 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 where 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 negative volatility premium, we assign all investors the benchmark parameter values in (43) and assume that the expected excess dividend growth and 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 (40) to compute the corresponding equilibrium expected return; but this time, as a guide to the intensity of trading, we also use (41) to compute $G(1)$, the probability of a trade within a year of purchase.

The top-right graph in Figure 6 plots the resulting relationship between expected return

[16] In our model, the risky assets are infinitely-lived. We have studied a variant of our model in which risky assets can 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 – earns a *higher* Sharpe ratio than does a long-horizon asset. More details are available on request.

and trade probability. It confirms that stocks with lower expected returns – stocks that are more "overpriced" – do indeed experience more turnover.

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 simpler, albeit related foundation, namely linear realization utility: in combination with a positive time discount rate, linear realization utility clearly 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, 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 will, on average, move higher 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 leads to 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. Li and Yang (2010) show that the link between realization utility and momentum *can* be formalized in an economy with both realization utility investors *and* expected utility investors. In their model, when realization utility investors sell a stock that is rising in value, the selling does depress the stock price because the demand from expected utility investors is not 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. However, we are careful to not only offer explanations for known facts, but to also suggest new predictions.

One set of predictions emerges from the graphs in Figure 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. Figure 5 suggests some other, more novel predictions, however: 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 the stock's average return and of his sensitivity to losses.

The prediction relating the probability of trade to a stock's average return is difficult to test because the average return perceived by individual investors may differ from the actual average return. Growth stocks, for example, have low average returns, but it is possible that some individual investors *perceive* them to have high average returns. The prediction relating the probability of trade to a stock's volatility, however, is much easier to test.

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. These 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 raises the possibility 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.

A second set of predictions builds on the idea that realization utility probably matters more for individual investors than for institutional investors: as trained professionals, institutional investors are less likely to think about their investing history as a series of investing episodes. With this in mind, we make the following predictions. First, since realization utility generates a strong disposition effect, the disposition effect should be more pronounced among individual investors than among institutional investors. Applying the same logic within the set of individual investors, we further predict that the disposition effect should be most pronounced for the least sophisticated individuals. Second, since realization utility generates risk-seeking, we should find that, controlling for other stock characteristics, individual investors overweight highly volatile stocks, while institutional investors underweight them.

Third, if realization utility generates Ang et al.'s (2006) negative volatility premium, then this premium should be stronger among stocks that are held and traded by individual investors. And fourth, if realization utility is in part responsible for the link between valuation and volume, this link should also be stronger among stocks held and traded by individuals.

Recently, researchers have begun to uncover evidence that speaks to some of these predictions. Frazzini (2006) confirms that the disposition effect is indeed stronger for individual investors than for institutional investors, while Dhar and Zhu (2006) show that it is most pronounced among the least sophisticated individuals. Han and Kumar (2008) find that individual investors overweight stocks with high idiosyncratic volatility, while institutional investors underweight them. They also report that Ang et al.'s (2006) negative volatility premium is indeed stronger among stocks that are heavily traded by individuals.

6 Conclusion

A number of authors have suggested that investors may derive utility from *realizing* gains and losses on assets that they own. We present a model of this "realization utility," derive its predictions, and show that it can shed light on a number of puzzling facts. While our analysis suggests that realization utility plays a role in actual financial markets, the only way to be sure is to test our model's predictions. We lay out some predictions in Section 5.3, but there are surely others that we have not reported there.

A useful feature of our model is that it allows the researcher to simulate long time series of stock prices and trading volume: given any stock price path, our model describes exactly how realization utility investors would trade along it. The properties of prices and volume in the simulated data can then be compared to the properties of prices and volume in actual data in order to see how well they match. This may be a fruitful way of testing for the importance of realization utility among actual investors.

7 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:

$$\begin{aligned}
& V(W_t, B_t) \\
= \; & \max\{u((1-k)W_t - B_t) + V((1-k)W_t, (1-k)W_t), \qquad (44) \\
& \quad v(D_{i,t})dt + (1-\rho dt)E_t[e^{-\delta dt}V(W_{t+dt}, B_{t+dt})] + \rho dt\,[u((1-k)W_t - B_t)]\} \\
= \; & \max\{u((1-k)W_t - B_t) + V((1-k)W_t, (1-k)W_t), \qquad (45) \\
& \quad v(D_{i,t}) + E_t\left[e^{-\delta dt}V(W_{t+dt}, B_{t+dt})\right] + \rho dt\,[u((1-k)W_t - B_t) - V(W_t, B_t)]\}.
\end{aligned}$$

The expression before the comma on the right-hand side of (44) shows what happens if 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 promptly invests in another stock. The expression after the comma on the right-hand side shows what happens if the investor instead holds his position for an infinitesimal period dt. First, 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 this interval and the investor's value function is simply the expected future value function, discounted back. With probability $1 - e^{-\rho dt} \approx \rho dt$, there *is* a liquidity shock and the investor sells his holdings and exits. This entails realization utility of $u((1-k)W_t - B_t)$.

We conjecture that the value function takes the form

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

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

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

Equation (46) implies that any solution to (10) must satisfy

$$U(g_t) \geq u((1-k)g_t - 1) + (1-k)g_t U(1) \qquad (47)$$

and

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

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 also continuously differentiable – this last condition is sometimes known as the "smooth pasting" condition. We then verify that $U(g_t)$ does indeed solve 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, then, equation (46) is maximized by the second term within the curly brackets and 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, equation (46) is maximized by the first term within the curly brackets and 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 equation (14) 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 $U(1) > 0$, 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 smooth around the liquidation point g_*. This implies

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

Solving these two equations, we obtain the expression for a in (15) and the following nonlinear equation for g_*:

$$(\gamma_1 - 1)\left(1 - \frac{k(\rho+\delta)(\rho + \frac{\alpha\beta}{1-k})}{\delta(\rho+\delta-\mu)}\right) g_*^{\gamma_1} - \frac{\gamma_1}{1-k} g_*^{\gamma_1 - 1} + 1 = 0. \tag{52}$$

Equation (52) has a unique solution in the range $(1, \infty)$. To see this, define

$$f(g) \equiv (\gamma_1 - 1)\left(1 - \frac{k(\rho+\delta)(\rho+\frac{\alpha\beta}{1-k})}{\delta(\rho+\delta-\mu)}\right) g^{\gamma_1} - \frac{\gamma_1}{1-k} g^{\gamma_1 - 1} + 1.$$

The parameter restriction in (11) implies that $\gamma_1 > 1$ and that $1 > \frac{k(\rho+\delta)(\rho+\frac{\alpha\beta}{1-k})}{\delta(\rho+\delta-\mu)}$. It is then straightforward to see that

$$f(1) < 0 \text{ and } f(\infty) > 0.$$

As a result, $f(g)$ has at least one root above 1. We now rule out the possibility that $f(g)$ has more than one root above 1. Suppose instead that $f(g)$ does have more than one root above 1. Then, it must have a local maximum $g_m > 1$ which satisfies $f'(g_m) = 0$ and $f''(g_m) < 0$. The condition $f'(g_m) = 0$ implies

$$g_m = \frac{1}{(1-k)\left(1 - \frac{k(\rho+\delta)(\rho+\frac{\alpha\beta}{1-k})}{\delta(\rho+\delta-\mu)}\right)}$$

and

$$f''(g_m) = \gamma_1(\gamma_1 - 1) g_m^{\gamma_1 - 3} \frac{1}{1-k} > 0.$$

The last inequality contradicts the initial assumption that g_m is a local maximum. The function $f(g)$ therefore has a unique root above 1.

We now verify that the constructed value function is indeed optimal. Substituting $V(W_t, B_t) = B_t U(g_t)$ into (10) and cancelling the B_t factor reduces the stopping problem to

$$\begin{aligned} U(g_t) \quad = \quad \max_{\tau \geq t} \quad & E_t\{\int_t^{\min\{\tau,\tau'\}} e^{-\delta(s-t)} \alpha\beta g_s ds \\ & + e^{-\delta(\tau-t)}[u((1-k)g_\tau - 1) + (1-k)g_\tau U(1)] I_{\{\tau<\tau'\}} \\ & + e^{-\delta(\tau'-t)} u((1-k)g_{\tau'} - 1) I_{\{\tau \geq \tau'\}}\}. \end{aligned} \tag{53}$$

We first verify that the function $U(g_t)$ summarized in equation (13) satisfies conditions (47) and (48). Define

$$f_1(g) \equiv (1-k)(1+U(1))g - 1.$$

Note that, by construction, $f_1(g)$ is a straight line which coincides with $U(g)$ for $g \geq g_*$. Since $\gamma_1 > 1$, $U(g)$ in equation (13) is a convex function. It must therefore lie above the straight line $f_1(g)$ for all $g < g_*$. Condition (47) is therefore satisfied.

We now check that condition (48) holds. Define

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

Note that for $g < g_*$, $H(g) = 0$ by construction. For $g \geq g_*$, $U(g) = f_1(g)$, 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 (15) into this expression, we obtain

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

The last equality follows by applying equation (16). Using (14), it is straightforward to show that if $\mu<\rho+\delta$, as assumed in parameter restriction (11), then $\rho+\delta-\mu\gamma_1>0$. Therefore, $H(g)<0$ for $g\ge g_*$. We have therefore confirmed that condition (48) holds for all $g_t\in(0,\infty)$.

Now note that $U(g)$ has an increasing derivative in $(0,g_*)$ and a derivative of $(1-k)(1+U(1))$ in (g_*,∞). $U'(g)$ is therefore bounded. Define the stopping time

$$\iota\equiv\min(\tau,\tau')\,,$$

where τ is *any* selling strategy and τ' is the time at which a liquidity shock arrives. Ito's lemma for twice-differentiable functions with absolutely continuous first derivatives – see, for example, Revuz and Yor (1999), Chapter 6 – implies

$$\begin{aligned} e^{-\delta(\iota-t)}U(g_\iota) &= U(g_t)+\int_t^\iota \sigma g_s U'(g_s)\,dZ_s \\ &+\int_t^\iota\left[\frac{1}{2}\sigma^2g_s^2U''(g_s)+\mu g_sU'(g_s)-(\rho+\delta)\,U(g_s)+\alpha\beta g_s+\rho u\left((1-k)g_s-1\right)\right]ds. \end{aligned}$$

The bound on $U'(g)$ implies that the first integral is a martingale, while condition (48) implies that the second integral is non-positive. We therefore have

$$U(g_t)\ge E_t\left[e^{-\delta(\iota-t)}U(g_\iota)\right]. \tag{54}$$

Note also, from condition (47), that

$$E_t\left[e^{-\delta(\iota-t)}U(g_\iota)\right]\ge E_t\left\{e^{-\delta(\iota-t)}\left[((1-k)g_\iota-1)+(1-k)g_\iota U(1)\right]\right\}. \tag{55}$$

Now consider the expression in the expectation operator of (53). If $\tau\le\tau'$, then $\iota=\tau$ and

$$\begin{aligned} & e^{-\delta(\tau-t)}[u((1-k)g_\tau-1)+(1-k)g_\tau U(1)]I_{\{\tau<\tau'\}}+e^{-\delta(\tau'-t)}u((1-k)g_{\tau'}-1)I_{\{\tau\ge\tau'\}} \\ = \;& e^{-\delta(\iota-t)}\left[((1-k)g_\iota-1)+(1-k)g_\iota U(1)\right]. \end{aligned} \tag{56}$$

If $\tau > \tau'$, so that $\iota = \tau'$, the expression satisfies

$$\begin{aligned}
& e^{-\delta(\tau-t)}[u((1-k)g_\tau - 1) + (1-k)g_\tau U(1)]I_{\{\tau<\tau'\}} + e^{-\delta(\tau'-t)}u((1-k)g_{\tau'} - 1)I_{\{\tau\geq\tau'\}} \\
\leq\ & e^{-\delta(\tau-t)}[u((1-k)g_\tau - 1) + (1-k)g_\tau U(1)]I_{\{\tau<\tau'\}} \\
& \quad + e^{-\delta(\tau'-t)}\left[u((1-k)g_{\tau'} - 1) + (1-k)g_{\tau'} U(1)\right] I_{\{\tau\geq\tau'\}} \\
=\ & e^{-\delta(\iota-t)}\left[((1-k)g_\iota - 1) + (1-k)g_\iota U(1)\right],
\end{aligned} \tag{57}$$

where the inequality follows from $U(1) \geq 0$. For any stopping time τ, we therefore have

$$\begin{aligned}
& E_t\{e^{-\delta(\tau-t)}[u((1-k)g_\tau - 1) + (1-k)g_\tau U(1)]I_{\{\tau<\tau'\}} + e^{-\delta(\tau'-t)}u((1-k)g_{\tau'} - 1)I_{\{\tau\geq\tau'\}}\} \\
\leq\ & E_t\left\{e^{-\delta(\iota-t)}\left[((1-k)g_\iota - 1) + (1-k)g_\iota U(1)\right]\right\} \\
\leq\ & E_t\left[e^{-\delta(\iota-t)} U(g_\iota)\right] \\
\leq\ & U(g_t),
\end{aligned}$$

where the first inequality follows from (56) and (57), the second from (55), and the third from (54). The constructed value function $U(g_t)$ is therefore at least as good as the value function generated by any alternative selling strategy. This completes the proof.

Proof of Proposition 2: We conjecture that the value function takes the form

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

Following the same logic as in the proof of Proposition 1, we find that $U(\cdot)$ again satisfies equation (46) and inequalities (47) and (48). The only difference is that $u(\cdot)$ now has the piecewise-linear form in (17).

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

$$\frac{1}{2}\sigma^2 g_t^2 U''(g_t) + \mu g_t U'(g_t) - (\rho + \delta) U(g_t) + \alpha\beta g_t + \rho u((1-k)g_t - 1) = 0. \tag{58}$$

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, $g_t = \frac{1}{1-k}$, is 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 (19), would imply

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

contradicting the desirable restriction that $U(g_t)$ be 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 $g_t \in \left(0, \frac{1}{1-k}\right)$, equation (58) 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

$$U\left(g_t\right)=bg_t^{\gamma_1}+\frac{\alpha\beta+\rho\lambda(1-k)}{\rho+\delta-\mu}g_t-\frac{\rho\lambda}{\rho+\delta}\quad\text{for}\quad g_t\in\left(0,\frac{1}{1-k}\right),\tag{59}$$

where γ_1 is defined in equation (14), and where b is determined below.

For $g_t\in\left(\frac{1}{1-k},g_*\right)$, equation (58) becomes

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

The solution to this equation is

$$U\left(g_t\right)=c_1g_t^{\gamma_1}+c_2g_t^{\gamma_2}+\frac{\alpha\beta+\rho(1-k)}{\rho+\delta-\mu}g_t-\frac{\rho}{\rho+\delta}\quad\text{for}\quad g_t\in\left(\frac{1}{1-k},g_*\right),$$

where

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

and where c_1 and c_2 are determined below.

The value function must be continuous and smooth around $g_t=\frac{1}{1-k}$. 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 equation (23), 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 equation (21), namely

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

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

$$U\left(g_t\right)=(1-k)g_t(1+U\left(1\right))-1.$$

The value function must be continuous and smooth around the liquidation point, so that

$$\begin{aligned}c_1g_*^{\gamma_1}+c_2g_*^{\gamma_2}+\frac{\alpha\beta+\rho\left(1-k\right)}{\rho+\delta-\mu}g_* &= (1-k)g_*(1+U\left(1\right))-\frac{\delta}{\rho+\delta}\\ c_1\gamma_1g_*^{\gamma_1-1}+c_2\gamma_2g_*^{\gamma_2-1}+\frac{\alpha\beta+\rho\left(1-k\right)}{\rho+\delta-\mu} &= (1-k)(1+U\left(1\right)).\end{aligned}$$

Since, from equation (59),

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

we obtain equation (24),

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

and equation (22),

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

All that remains is to verify the optimality of the constructed value function. This part of the derivation is similar to the final part of the proof of Proposition 1. For space reasons, we do not repeat it here.

Proof of Proposition 3: We solve the decision problem in (31) using a technique very similar to 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 capital gain, standard deviation, and transaction cost of stock i, respectively. We also note that $U(1) = 0$ in equilibrium. It is then straightforward to obtain the results in Proposition 3.

Proof of Proposition 4: Define

$$x_t \equiv \ln(g_t) \quad \text{and} \quad x_* \equiv \ln(g_*).$$

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 level 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 can 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.

8 References

Ang, A., Hodrick, R., Xing, Y., and X. Zhang (2006), "The Cross-Section of Volatility and Expected Returns," *Journal of Finance* 61, 259-299.

Baker, M. and J. Wurgler (2007), "Investor Sentiment in the Stock Market," *Journal of Economic Perspectives* 21, 129-151.

Baker, M., Pan, X., and J. Wurgler (2009), "A Reference Point Theory of Mergers and Acquisitions," Working paper, New York University.

Barber, B., Lee, Y., Liu, Y., and T. Odean (2009), "Just How Much Do Investors Lose From Trade?" *Review of Financial Studies* 22, 151-186.

Barber, B. and T. Odean (2000), "Trading is Hazardous to Your Wealth: The Common Stock Performance of Individual Investors," *Journal of Finance* 55, 773-806.

Barberis, N. and M. Huang (2001), "Mental Accounting, Loss Aversion, and Individual Stock Returns," *Journal of Finance* 56, 1247-1292.

Barberis, N., Huang, M., and T. Santos (2001), "Prospect Theory and Asset Prices," *Quarterly Journal of Economics* 116, 1-53.

Barberis, N. and W. Xiong (2009), "What Drives the Disposition Effect? An Analysis of a Long-standing Preference-based Explanation," *Journal of Finance* 64, 751-784.

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

Dhar, R. and N. Zhu (2006), "Up Close and Personal: Investor Sophistication and the Disposition Effect," *Management Science* 52, 726-740.

Frazzini, A. (2006), "The Disposition Effect and Underreaction to News," *Journal of Finance* 61, 2017-2046.

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

Griffin, J., Nardari, F., and R. Stulz (2007), "Do Investors Trade More When Stocks Have Performed Well? Evidence from 46 Countries," *Review of Financial Studies* 20, 905-951.

Grinblatt, M. and B. Han (2005), "Prospect Theory, Mental Accounting, and Momentum," *Journal of Financial Economics* 78, 311-339.

Grinblatt, M. and M. Keloharju (2001), "What Makes Investors Trade?" *Journal of Finance* 56, 589-616.

Grinblatt, M. and M. Keloharju (2009), "Sensation-seeking, Overconfidence, and Trading Activity," *Journal of Finance* 64, 549-578.

Han, B. and A. Kumar (2008), "Retail Clienteles and the Idiosyncratic Volatility Puzzle," Working paper, University of Texas at Austin.

Heath, C., Huddart, S., and M. Lang (1999), "Psychological Factors and Stock Option Exercise," *Quarterly Journal of Economics* 114, 601-627.

Hong, H. and J. Stein (2007), "Disagreement and the Stock Market," *Journal of Economic Perspectives* 12, 109-128.

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

Li, Y. and L. Yang (2010), "Prospect Theory and Price Momentum," Working paper, University of Toronto.

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.

Revuz, D. and M. Yor (1999), *Continuous Martingales and Brownian Motion*, Springer: New York.

Shefrin, H. and M. Statman (1985), "The Disposition to Sell Winners too Early and Ride Losers too Long," *Journal of Finance* 40, 777-790.

Statman, M., Thorley, S., and K. Vorkink (2006), "Investor Overconfidence and Trading Volume," *Review of Financial Studies* 19, 1531-1565.

Thaler, R. (1999), "Mental Accounting Matters," *Journal of Behavioral Decision-making* 12, 183-206.

Weber, M. and C. Camerer (1998), "The Disposition Effect in Securities Trading: An Experimental Analysis," *Journal of Economic Behavior and Organization* 33, 167-184.

Zacks, J. and K. Swallow (2007), "Event Segmentation," *Current Directions in Psychological Science* 16, 80-84.

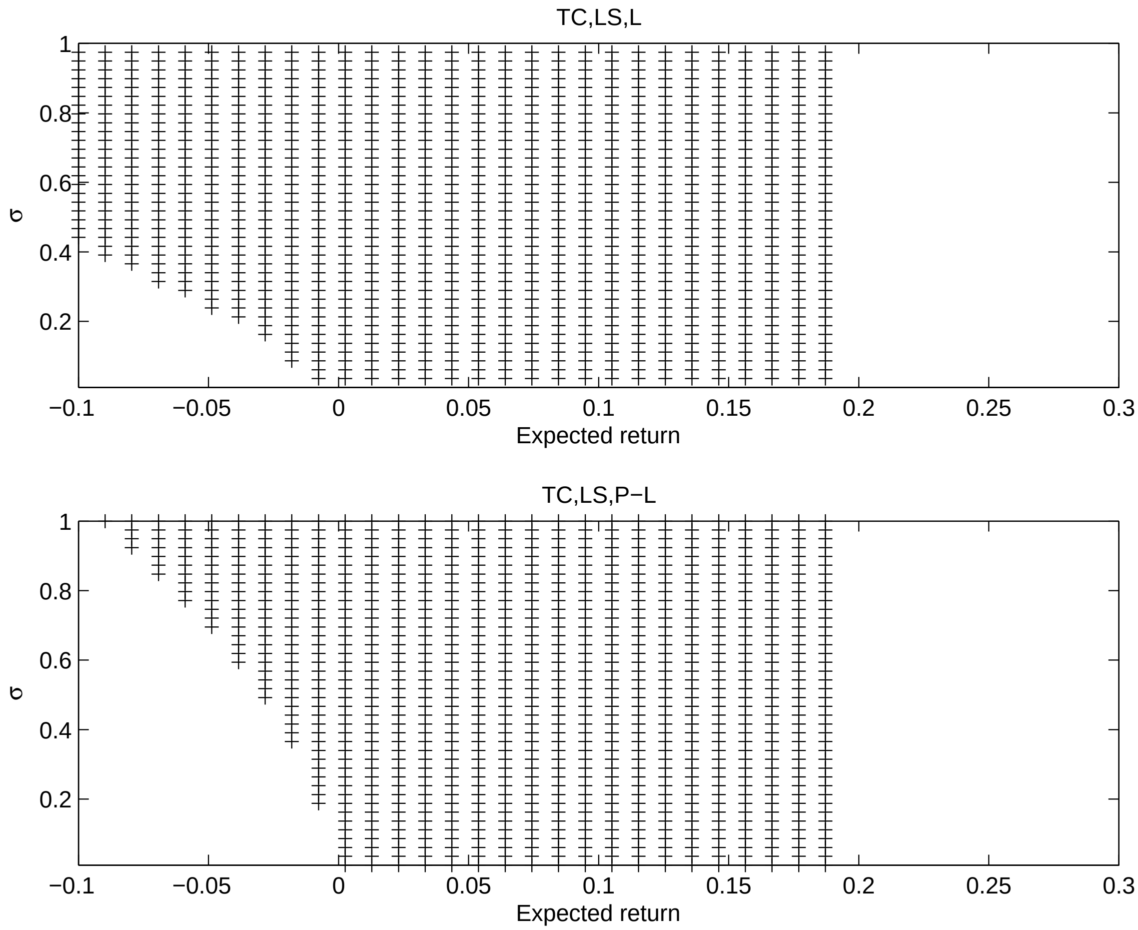

Figure 1. The graphs show, for an investor who derives utility from realized gains and losses, the range of values of a stock's expected excess return and standard deviation σ for which the investor is willing both to buy the stock and to sell it once its price reaches a sufficiently high liquidation point. The top graph corresponds to a model that allows for a transaction cost (TC) and an exogeneous liquidity shock (LS), and in which realization utility has a linear functional form (L). The bottom graph corresponds to a model that also allows for a transaction cost and an exogeneous liquidity shock, but in which realization utility has a piecewise-linear functional form (P-L), so that the investor is 1.5 times as sensitive to realized losses as to realized gains.

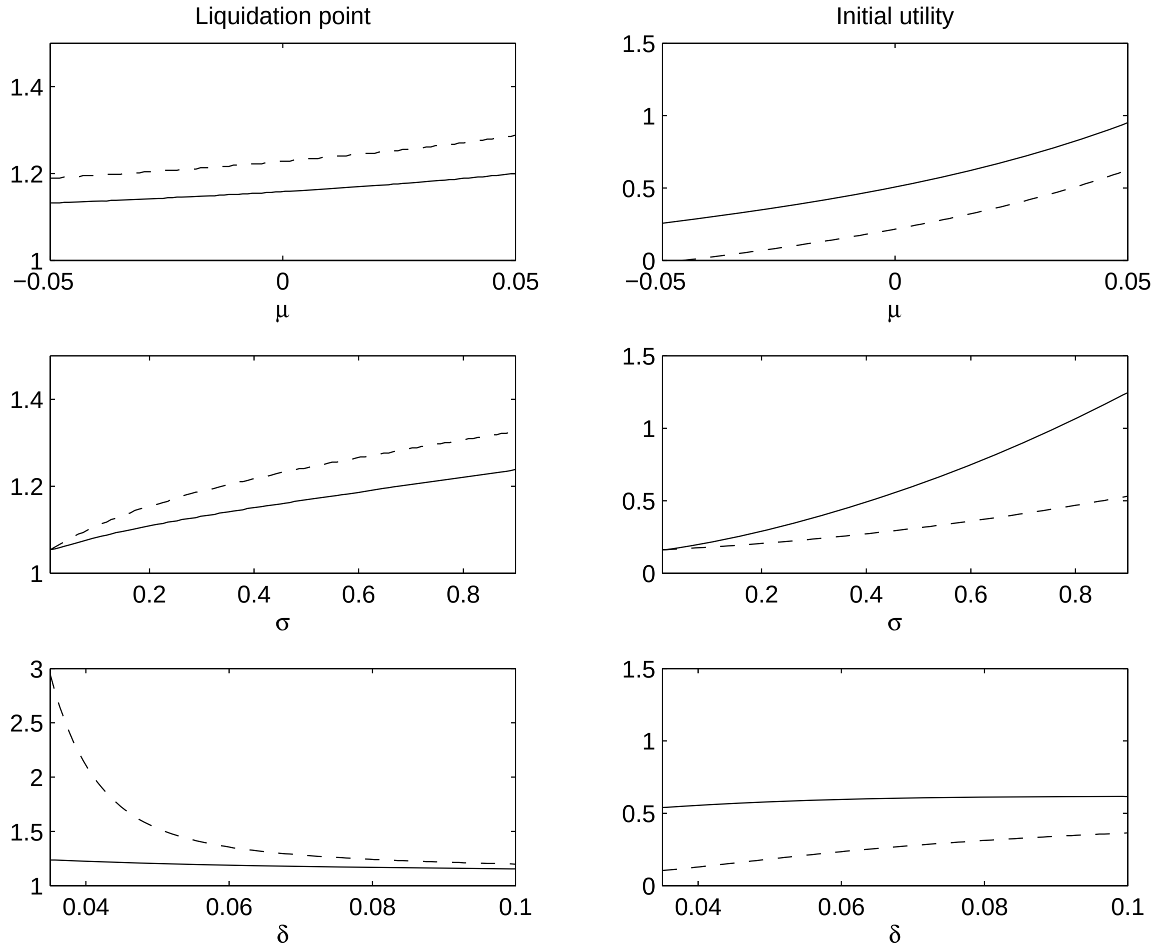

Figure 2. The graphs show, for an investor who derives utility from realized gains and losses, how the liquidation point at which he sells a stock and the initial utility from buying it depend on the stock's expected excess capital gain μ, its standard deviation σ, and the time discount rate δ. The solid lines correspond to a model that allows for a transaction cost and an exogeneous liquidity shock, and in which realization utility has a linear functional form. The dashed lines correspond to a model that also allows for a transaction cost and an exogeneous liquidity shock, but 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.

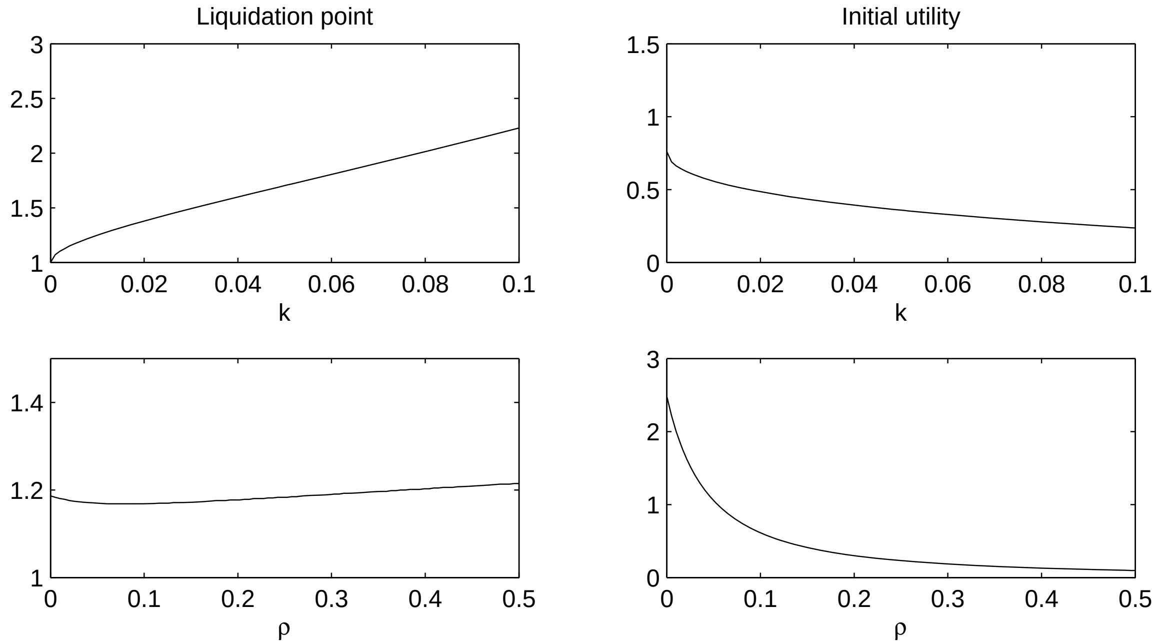

Figure 3. The graphs show, for an investor who derives utility from realized gains and losses, how the liquidation point at which he sells a stock and the initial utility from buying it depend on the transaction cost k and the arrival rate ρ of an exogeneous liquidity shock. In these computations, realization utility has a linear functional form.

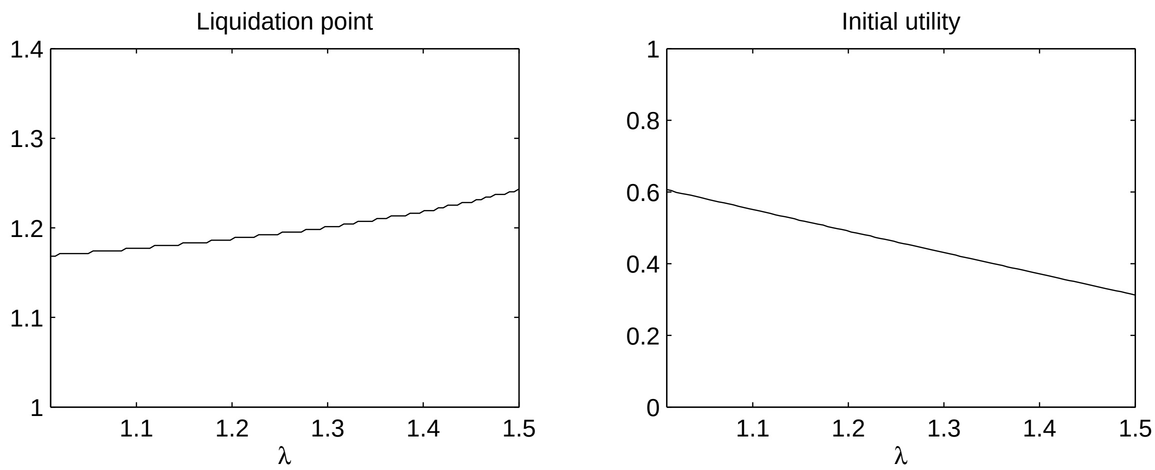

Figure 4. The graphs show, for an investor who derives utility from realized gains and losses, how the liquidation point at which he sells a stock and the initial utility from buying it depend on λ, his relative sensitivity to realized losses as opposed to realized gains. The computations are based on a model that allows for a transaction cost and an exogeneous liquidity shock.

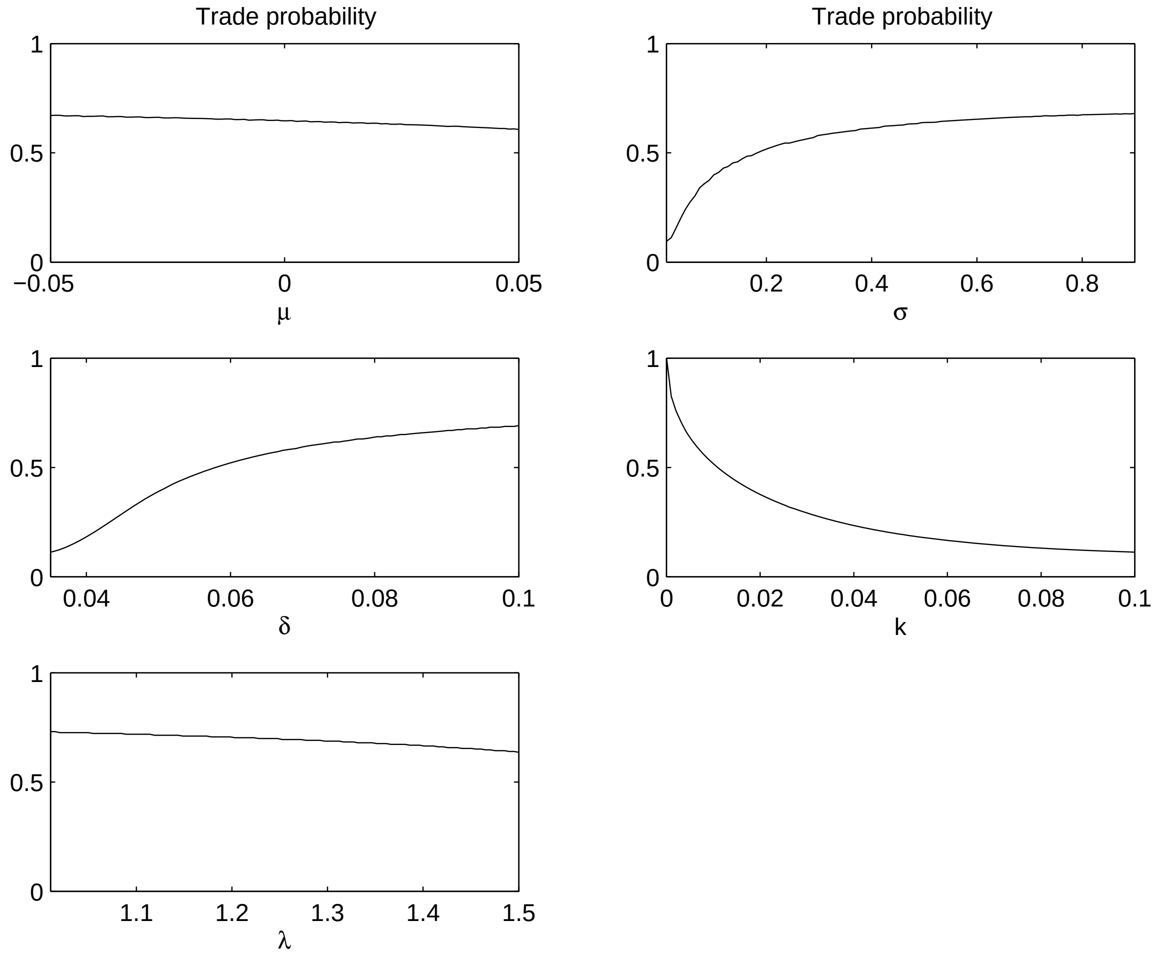

Figure 5. The graphs show, for an investor who derives utility from realized gains and losses, how the probability that the investor will sell a specific stock within a year of buying it depends on the stock's expected excess capital gain μ, its standard deviation σ, the time discount rate δ, the transaction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

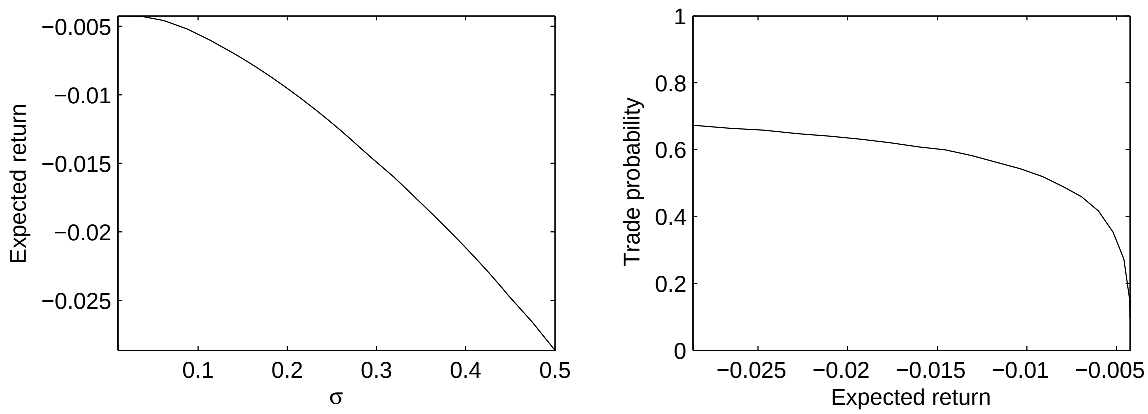

Figure 6. The top-left graph shows, for an economy populated by investors who derive utility from realized gains and losses, the equilibrium relationship between expected excess return and standard deviation in a cross-section of stocks. The top-right graph shows, for the same cross-section, the equilibrium relationship between expected excess return and trading intensity. The computations are based on a model that allows for a transaction cost and an exogeneous liquidity shock, and in which realization utility has a piecewise-linear functional form, so that investors are 1.5 times more sensitive to realized losses than to realized gains.