

NBER WORKING PAPER SERIES

REALIZATION UTILITY

Nicholas C. Barberis
Wei Xiong

Working Paper 14440
<http://www.nber.org/papers/w14440>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2008

We thank Bige Kahraman for able research assistance. We are also grateful to Daniel Benjamin, Patrick Bolton, Erik Eyster, Antonio Mele, Chris Rogers, Paul Tetlock, Dimitri Vayanos, 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, the Gerzensee Summer Symposium, and the NBER for helpful comments. Comments are welcome at nick.barberis@yale.edu and wxiong@princeton.edu. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2008 by Nicholas C. Barberis and Wei Xiong. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Realization Utility
Nicholas C. Barberis and Wei Xiong
NBER Working Paper No. 14440
October 2008
JEL No. G11,G12

ABSTRACT

We study the possibility that, aside from standard sources of utility, investors also derive utility from realizing gains and losses on assets that they own. We propose a tractable 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 negative premium to volatility in the cross-section, and the heavy trading of highly valued assets. Underlying some of these applications is one of our model's more novel predictions: that, even if the form of realization utility is linear or concave, investors can be risk-seeking.

Nicholas C. Barberis
Yale School of Management
135 Prospect Street
P O Box 208200
New Haven, CT 06520-8200
and NBER
nick.barberis@yale.edu

Wei Xiong
Princeton University
Department of Economics
Bendheim Center for Finance
Princeton, NJ 08450
and NBER
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 also derive utility from another source, namely from *realized gains and losses* on assets that they own. Suppose, for example, that an investor buys a stock and then, a few months later, sells it. We analyze a model in which the investor gets a jolt of utility right then, at the moment of sale, and where the utility term depends on the size of the gain or loss realized – utility is positive if the stock is sold at a gain relative to purchase price and negative otherwise. We label this source of utility “realization utility.”

Why might an investor derive utility from realizing a gain or loss? We think that realization utility is primarily a consequence of what Thaler (1999) calls “mental accounting,” in other words, a consequence of the way people organize and evaluate their financial investments. Under this view, when people think about their investments, they do not do so purely in terms of overall portfolio return. Rather, they often think about their investing history 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.

Realization utility is a natural consequence of mental accounting. When an investor sells a stock at a gain, he feels good because he is creating a *positive* investing episode, one that he can look back on and talk about with pleasure. Conversely, if he sells a stock at a loss, he feels bad because he is creating a *negative* investing episode, one that will be painful to look back on and talk about.

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 track their performance using overall portfolio return rather than stock-level mental accounting. We might 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, particularly when thought about in terms of mental accounting. 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 must construct a model of

realization utility and then check to see if the model explains a range of facts and leads to new predictions.

In this paper, we take up this challenge. We develop a tractable 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 surprisingly wide range of applications. Finally, we lay out some testable new predictions. We start with a partial equilibrium framework but also show how realization utility can be embedded in a full equilibrium model. As a result, we are able 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 random liquidity shocks: 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 the case of piecewise-linear utility.

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 stock returns, the time discount rate, the level of transaction costs, and the frequency of liquidity shocks. The model also allows us to compute the probability that, within any given interval after first buying a stock, the investor sells it. We look at how this probability – a measure of trading frequency – 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 experience a large drop 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. The model also predicts that more volatile stocks will be traded more frequently: roughly speaking, a more volatile stock reaches its liquidation point more rapidly.

We link our model to a wide range of financial phenomena. In particular, we show that it can shed light on the subpar trading performance of individual investors (Odean, 1999; Barber and Odean, 2000), 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 negative volatility premium (Ang et al., 2006), and the heavy trading associated with highly valued assets – as, for example, in the technology sector in the late 1990s (Hong and Stein, 2007).

To give the reader an early sense of some of the ideas behind these applications, we briefly pick out two. Our model predicts that individual investors – the investor group that is more likely to experience realization utility – will have a much greater propensity to sell a stock once it moves above its historical high. Imagine a stock which 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 of \$45 or higher 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 the propensity to sell.

Our model also explains the heavy trading that is often associated with highly valued assets. To see this, note that, in an economy where many investors care about realization utility, more volatile stocks will be both more heavily traded – such stocks reach their liquidation points faster – *and* more highly valued: since realization utility investors like volatility, they will collectively push the prices of volatile stocks up. The equilibrium version of our model therefore predicts a coincidence of high valuations and heavy trading; and moreover, that this phenomenon will occur for assets whose value is particularly uncertain. Under this view, the late 1990s were years in which 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, generating high trading volume.

As noted above, realization utility offers a simple way of understanding a range of financial phenomena. In most cases, we did not foresee the link between realization utility and the applications we describe: we saw the link only after completing our analysis. Nonetheless, we are careful to not only offer explanations for known facts, but to also suggest new predictions. We discuss a number of these new predictions in the final section of the paper.

An early discussion of realization utility can be found in Shefrin and Statman (1985). They propose it, in combination with prospect theory, as a way of understanding the disposition effect, and present a two-period numerical example. More recently, Barberis and

Xiong (2008) briefly analyze a two-period model of realization utility, again in combination with prospect theory and again with the disposition effect as the eventual application.

In this paper, we present the first comprehensive analysis of realization utility. We move beyond the two-period setting and work in an infinite horizon framework. We allow for realistic features of trading such as transaction costs and liquidity shocks. We analyze the investor's trading strategy along several dimensions, including trading frequency. We study realization utility not only in partial equilibrium, but also in a full equilibrium. And we consider a wide range of applications, of which the disposition effect is just one.

In Section 2, we present the baseline model, one that assumes linear realization utility. In Section 3, we consider an alternative preference specification, namely piecewise-linear utility. In Section 4, we show how realization utility can be embedded in a full equilibrium framework. Section 5 discusses a range of applications and new predictions, and Section 6 concludes.

2 A Model of Realization Utility

In this paper, we present an analysis of realization utility. Earlier papers, such as Barberis and Huang (2001), have studied models in which investors derive utility from fluctuations in the value of specific assets that they are holding. In these models, investors derive utility from a gain in the value of a stock, whether that gain is realized or not. Here, we take a different approach. We posit that, from the perspective of preferences, there is a distinction between a paper gain and a realized gain, and that realized gains and losses give rise to their own separate component of utility.

Before presenting the model, we briefly note two of our assumptions. First, we assume that realization utility is defined at the level of an individual asset rather than at the level of a portfolio. When an investor makes a sale, he sells a specific asset. Since realization utility is associated with the act of selling, it makes sense that it be defined at the asset level. This assumption has little bite in our baseline model because, in this model, the investor holds at most one stock at any time. However, it becomes more important when we discuss an extension of our model in which the investor can hold several stocks simultaneously.

A second assumption concerns the functional form for realization utility. In this section, we focus on the simplest possibility, 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 the case of piecewise-linear utility.

Consider an investor who 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 net return of zero, and N risky assets indexed by $i \in \{1, \dots, N\}$. The price of risky asset i , $S_{i,t}$, follows

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

where $dZ_{i,t}$ is a Brownian motion and where, for $i \neq j$, $dZ_{i,t}$ and $dZ_{j,t}$ may be correlated. For now, we take μ and σ to be the same for all the risky assets. Perhaps 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.¹

For simplicity, 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: 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

$$\frac{dW_t}{W_t} = \sum_{i=1}^N (\mu dt + \sigma dZ_{i,t}) \theta_{i,t} - kl_t, \quad (2)$$

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 $j \neq i$.

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

$$B_t = W_s, \text{ where } s = \min\{\tau \in [0, t] : \theta_{i,\tau'} = 1 \text{ for all } \tau' \in [\tau, t]\}. \quad (3)$$

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

$$u((1 - k)W_t - B_t). \quad (4)$$

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,

$$u(x) = x. \quad (5)$$

¹We use a continuous-time framework because this allows us to solve the model analytically. We have also studied the discrete-time analog of our model. The results are similar, but can only be obtained numerically.

We emphasize that the investor only receives the burst of utility in (4) 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.²

The investor also faces the possibility of a random liquidity shock whose arrival is governed by a Poisson process with parameter ρ . When a shock occurs, the investor immediately sells his holdings and exits the asset markets. We think of this shock as an unexpected consumption need which forces the investor to draw on the funds in his brokerage account. Our results do not depend on the presence of a liquidity shock, but the shock is useful nonetheless: it ensures that the investor cares not only about realized gains and losses, but also about *paper* gains and losses. After all, even if a real-life investor cares about realization utility, he almost certainly also derives utility from paper gains and losses.

Suppose that, at time t , the investor's wealth is allocated to a stock. The investor's value function is a function of the current asset value, W_t , and of the asset's cost basis, B_t , and we denote it as $V(W_t, B_t)$. Since the utility function in (5) is 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 $\zeta > 0$,

$$V(\zeta W_t, \zeta B_t) = \zeta V(W_t, B_t). \quad (6)$$

Now suppose that, for some positive W ,

$$V(W, W) > 0. \quad (7)$$

Note that $V(W, W)$ is the value function from investing an amount W in a stock *now*, so that the asset's current value and cost basis are both equal to W . Since $V(W_t, B_t)$ is homogeneous of degree one, if condition (7) holds for some positive W , then it holds for all positive W . Given a positive time discount rate, condition (7) then immediately 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 (7) implies that, if, at any time $t > 0$, the investor sells a position in a stock, he will then immediately put the proceeds into another stock. We will later present a figure showing the range of parameter values for which condition (7) holds.

²The investor only receives realization utility when he sells stock, not when he sells the risk-free asset. The reason is that, since the risk-free rate is zero, the realized gain or loss from selling the risk-free asset is always zero. We also assume that the investor does not incur a transaction cost when selling the risk-free asset.

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

$$V(W_t, B_t) = \max_{\tau \geq t} E_t \{ e^{-\delta(\tau-t)} [u((1-k)W_\tau - B_\tau) + V((1-k)W_\tau, (1-k)W_\tau)] I_{\{\tau < \tau'\}} + e^{-\delta(\tau'-t)} u((1-k)W_{\tau'} - B_{\tau'}) I_{\{\tau \geq \tau'\}} \}, \quad (8)$$

subject to equations (2), (3), and (5). 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:

$$\max \left\{ 1, \frac{\delta}{\delta - \rho k} \right\} \mu < \rho + \delta. \quad (9)$$

To understand the formulation in (8), 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 current position. Suppose first that $\tau < \tau'$, so that the investor voluntarily sells the stock before a liquidity shock arrives. In this case, only the terms within the square parentheses are non-zero: when he liquidates his position at time τ , the investor receives a burst of utility $u((1-k)W_\tau - B_\tau)$ 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. In this case, only the final term is non-zero: the investor receives realization utility $u((1-k)W_{\tau'} - B_{\tau'})$ from the gain or loss at the moment of exit.

The proposition below summarizes the solution to the decision problem in (8). Put simply, if the investor buys a stock, he voluntarily sells it only if its price rises a sufficient amount above the purchase price. The variable

$$g_t = \frac{W_t}{B_t} \quad (10)$$

– in words, the percentage change in value, since purchase, 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 (8) 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

$$U(g_t) = \begin{cases} ag_t^{\gamma_1} + \frac{\rho(1-k)}{\rho+\delta-\mu}g_t - \frac{\rho}{\rho+\delta} & \text{if } g_t < g_* \\ (1-k)g_t(1+U(1)) - 1 & \text{if } g_t \geq g_* \end{cases}, \quad (11)$$

where

$$\gamma_1 = \frac{1}{\sigma^2} \left[\sqrt{\left(\mu - \frac{1}{2}\sigma^2\right)^2 + 2(\rho + \delta)\sigma^2} - \left(\mu - \frac{1}{2}\sigma^2\right) \right] > 0 \quad (12)$$

and

$$a = \frac{\delta}{g_*^{\gamma_1}(\gamma_1 - 1)(\rho + \delta)}. \quad (13)$$

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

$$(\gamma_1 - 1) \left(1 - \frac{\rho k(\rho + \delta)}{\delta(\rho + \delta - \mu)} \right) g_*^{\gamma_1} - \frac{\gamma_1}{1 - k} g_*^{\gamma_1 - 1} + 1 = 0. \quad (14)$$

We prove the proposition in the Appendix. In brief, 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.³

In summary, then, if the model parameters are such that initial utility per unit wealth $U(1)$ is positive, 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 $U(1) > 0$ does not hold, the investor instead allocates his wealth to the risk-free asset forever.

For expositional simplicity, we have assumed that the investor holds at most one stock at any time. However, the solution to (8) can also tell us how the investor trades in a setting where he holds *several* stocks at the same time. Suppose that he starts 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 stock market. Under these assumptions, the investor's decision problem is "separable" across the different stocks he is holding and the solution to (8) describes how he trades each one of his stocks.⁴

Results

³Since $g_* \geq 1$, the term $U(1)$ which appears in the second row of equation (11) can be directly obtained from the first row of equation (11). Specifically, it equals $a + \rho(1 - k)/(\rho + \delta - \mu) - \rho/(\rho + \delta)$.

⁴A corollary to Proposition 1 – one that holds even for the piecewise-linear specification we consider in Section 3 – is that, under 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 $W_0U(1)/2 + W_0U(1)/2$.

In this section, and again in Section 3, we draw out the implications of our model through two kinds of analysis. First, we look at the range of values of the model parameters for which the investor is willing to buy a stock at time 0. Second, we look at how the liquidation point g_* and initial utility $U(1)$ 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 investors.

The shaded area in the top graph in Figure 1 shows the range of values of the expected stock return μ and standard deviation of stock returns σ that satisfy $U(1) \geq 0$ – so that the investor is willing to buy a stock at time 0 – but also the restriction in (9), so that the investor is willing to sell the stock at a finite liquidation point. To create the graph, we need to assign values to the three remaining parameters, δ , k , and ρ . We set the time discount rate to $\delta = 0.08$ and the transaction cost to $k = 0.01$, which is of a similar order of magnitude to the transaction cost estimated by Barber and Odean (2000) for discount brokerage customers. Finally, we set $\rho = 0.1$. The probability of a liquidity shock over the course of a year is therefore $1 - e^{-0.1} \approx 0.1$.

The graph illustrates an interesting feature of our model, namely that the investor is willing to invest in 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.

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

To construct the graphs, we start with a set of benchmark parameter values. We use the same benchmark values throughout the paper. We set the expected excess stock return to $\mu = 0.03$ and the standard deviation of stock returns to $\sigma = 0.5$. We use a time discount rate of $\delta = 0.08$, a transaction cost of $k = 0.01$, and a liquidity shock intensity of $\rho = 0.1$. The graphs in Figures 2 and 3 show what happens as we vary each of μ , σ , δ , k , and ρ in turn, keeping the other parameters fixed at their benchmark values.

The top graphs in Figure 2 show that, as we would expect, initial utility is increasing in the mean stock return μ . 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 one of the important implications of our model: that, as stock return volatility goes up, the investor’s 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 (11) is convex.⁵ A similar intuition explains why, in the middle-left graph, the liquidation point is increasing in volatility.

Our model’s risk-seeking implication assumes that the investor is “sophisticated”: he is aware of how he will experience realization utility in the future. It is hard to know, based on introspection alone, whether actual investors are sophisticated in this way. We note two things about this. First, the risk-seeking result only requires *partial* sophistication, not full sophistication. Second, precisely because it is hard to introspect about how the typical investor thinks, we prefer to take another approach: to simply lay out the predictions of sophistication and then check if these predictions shed light on puzzling facts. In Section 5, we discuss some evidence that is consistent with at least partial, if not full sophistication, on the part of actual investors.

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 cannot wait very long before realizing a gain.

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.

⁵The parameter restriction in (9) implies $\gamma_1 > 1$ and $a > 0$, which, in turn, implies the convexity of $U(\cdot)$.

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 the investor's 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 confirm that, when $k = 0$, equation (14) 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.

We suspect that, in reality, the transaction cost is not the only thing that deters an investor who cares about realization utility from realizing all gains immediately. For example, in reality, an investor may be reluctant to sell a stock until he has located *another* attractive stock to transfer the proceeds of a sale to. While our model assumes that an alternative investment is always available, it may, in reality, take the investor some time to locate such an investment. The harder it is to find an alternative attractive investment, the more reluctant the investor will be to liquidate his current position.

To confirm this idea, we have analyzed an extension of our model in which, after selling his position in a stock, the investor is required to keep the proceeds in the risk-free asset for T periods, during which time he searches for a new stock to buy; the higher T is, the harder it is to uncover a new investment opportunity. We find that the investor's optimal strategy in this model is, once again, to sell a position in stock only if the stock price rises sufficiently far above the purchase price. In line with the intuition above, we find that the liquidation point is indeed increasing in T : the harder it is to find a new investment opportunity, the more reluctant the investor is to liquidate his current position.⁶

The bottom graphs in Figure 3 show how the liquidation point g_* and initial utility $U(1)$ depend on the intensity of the liquidity shock ρ . The bottom-left graph shows that the liquidation point depends on ρ in a non-monotonic way. There are two factors 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. In the presence of liquidity shocks, 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, there is a second factor which 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

⁶The details of this analysis are available on request.

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, the investor's utility falls. Since a liquidity shock can force the investor to exit the stock market with a painful loss, it lowers his utility.

3 An Alternative Specification: Piecewise-linear Utility

In Section 2, we took the functional form for realization utility $u(\cdot)$ to be linear. We did this so as to show that we do not need strong assumptions about $u(\cdot)$ in order to draw interesting implications out of realization utility. We now look at what happens when $u(\cdot)$ is piecewise-linear, rather than linear, so that the investor is more sensitive to realized losses than to realized gains:

$$u(x) = \begin{cases} x & \text{if } x \geq 0 \\ \lambda x & \text{if } x < 0 \end{cases}, \quad \lambda > 1. \quad (15)$$

The parameter λ controls the relative sensitivity to realized losses as opposed to realized gains.⁷

The investor's decision problem is now

$$\begin{aligned} V(W_t, B_t) = \max_{\tau \geq t} & E_t \{ e^{-\delta(\tau-t)} [u((1-k)W_\tau - B_\tau) + V((1-k)W_\tau, (1-k)W_\tau)] I_{\{\tau < \tau'\}} \\ & + e^{-\delta(\tau'-t)} u((1-k)W_{\tau'} - B_{\tau'}) I_{\{\tau \geq \tau'\}} \}, \end{aligned} \quad (16)$$

subject to (2), (3), and (15). This is the same as decision problem (8) in Section 2 except that $u(\cdot)$ is no longer linear, but instead takes the form in (15).

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 (16) will sell a position in stock once the gain $g_t = W_t/B_t$ reaches

⁷It 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 as opposed to paper gains, we should not conclude that they are necessarily also more sensitive to realized losses as opposed to realized gains.

a liquidation point $g_t = g_* \geq 1$; if the transaction cost k is positive, $g_* > 1$. The investor's value function is $V(W_t, B_t) = B_t U(g_t)$, where

$$U(g_t) = \begin{cases} b g_t^{\gamma_1} + \frac{\rho \lambda (1-k)}{\rho + \delta - \mu} g_t - \frac{\rho \lambda}{\rho + \delta} & \text{if } g_t \in \left(0, \frac{1}{1-k}\right) \\ c_1 g_t^{\gamma_1} + c_2 g_t^{\gamma_2} + \frac{\rho(1-k)}{\rho + \delta - \mu} g_t - \frac{\rho}{\rho + \delta} & \text{if } g_t \in \left(\frac{1}{1-k}, g_*\right) \\ (1-k)g_t(1 + U(1)) - 1 & \text{if } g_t \in (g_*, \infty) \end{cases}, \quad (17)$$

where γ_1 is defined in equation (12), where

$$\gamma_2 = -\frac{1}{\sigma^2} \left[\sqrt{\left(\mu - \frac{1}{2}\sigma^2\right)^2 + 2(\rho + \delta)\sigma^2} + \left(\mu - \frac{1}{2}\sigma^2\right) \right] < 0, \quad (18)$$

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

$$c_2 = \frac{(\lambda - 1)\rho(1-k)^{\gamma_2}(\mu\gamma_1 - \rho - \delta)}{(\gamma_1 - \gamma_2)(\rho + \delta - \mu)(\rho + \delta)} \quad (19)$$

$$(\gamma_1 - 1)c_1 g_*^{\gamma_1} + (\gamma_2 - 1)c_2 g_*^{\gamma_2} = \frac{\delta}{\rho + \delta} \quad (20)$$

$$c_1 \left(\frac{1}{1-k}\right)^{\gamma_1} + c_2 \left(\frac{1}{1-k}\right)^{\gamma_2} = b \left(\frac{1}{1-k}\right)^{\gamma_1} + \frac{(\lambda - 1)\mu\rho}{(\rho + \delta - \mu)(\rho + \delta)} \quad (21)$$

$$c_1 g_*^{\gamma_1} + c_2 g_*^{\gamma_2} + \frac{(1-k)(\mu - \delta)}{\rho + \delta - \mu} g_* + \frac{\delta}{\rho + \delta} = (1-k)g_* \left(b + \frac{\rho\lambda(\mu - k\rho - k\delta)}{(\rho + \delta)(\rho + \delta - \mu)} \right). \quad (22)$$

Specifically, given values for μ , σ , δ , k , ρ , and λ , we first use equation (19) to find c_2 ; we then obtain c_1 from equation (20); we then use equation (21) to find b ; finally, equation (22) 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 μ and σ for which the investor is both willing to buy stock at time 0, so that $U(1)$, from (17), is positive, and also to sell it at a finite liquidation point. We set δ , k , and ρ to the benchmark values from before, namely 0.08, 0.01, and 0.1, respectively. We further 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 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 will therefore only invest 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 which he can enjoy realizing.

The graphs in Figure 4 show how the liquidation point g_* and initial utility $U(1)$ depend on the sensitivity to losses λ . These graphs vary λ while maintaining

$$(\mu, \sigma, \delta, k, \rho) = (0.03, 0.5, 0.08, 0.01, 0.1). \quad (23)$$

In the top-left graph, we see that the more sensitive the investor is to losses, the higher the liquidation point: if the investor is holding a stock with a gain, he is reluctant to realize that 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 top-right graph shows that, as the sensitivity to losses goes up, the investor's 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,

$$(\mu, \sigma, \delta, k, \rho, \lambda) = (0.03, 0.5, 0.08, 0.01, 0.1, 1.5). \quad (24)$$

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. As expected from Figure 4, a higher value of λ means a higher liquidation point g_* and a lower initial utility $U(1)$.

The dashed line in the middle-right graph of Figure 2 deserves particular attention. It shows that, for the benchmark values in (24), the investor's initial utility $U(1)$ is still increasing in stock volatility σ . Put differently, even though the form of 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 will reverse, so that $U(1)$ becomes a decreasing function of σ .

4 Asset Pricing

In Sections 2 and 3, we studied realization utility in a partial equilibrium context. In this section, we show how it can be embedded in a full equilibrium. This, in turn, 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, ex-ante, 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 or if it leads to new predictions which can be tested and confirmed.

In general, embedding non-standard preferences in a full equilibrium can be challenging. To make headway, we therefore 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 $i \in \{1, \dots, N\}$. The risk-free asset is in perfectly elastic supply and earns a net return of zero. The risky stocks are in limited supply and the price process for stock i is

$$\frac{dS_{i,t}}{S_{i,t}} = \mu_i dt + \sigma_i dZ_{i,t}, \quad (25)$$

where $dZ_{i,t}$ is a Brownian motion and where, for $i \neq j$, $dZ_{i,t}$ and $dZ_{j,t}$ may be correlated. The parameter σ_i is constant over time. We assume, for now, that μ_i is also constant over time, and confirm this assumption later. Note that, in contrast to Sections 2 and 3, the N stocks can now differ in their expected returns and standard deviations.

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 , and λ 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

$$V_i(W, W) = 0, \quad i = 1, \dots, N, \quad (26)$$

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 (26) the appropriate equilibrium conditions? Note that, under the conditions in (26), 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 (26) mean that they are happy to then be assigned to the risk-free asset. Finally, the conditions in (26) mean that, if some investors do sell their holdings of stock i , we can reassign other investors from the risk-free asset to stock i , thereby again clearing the market in stock i .

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

$$V_i(W_t, B_t) = \max_{\tau \geq t} E_t \{ e^{-\delta(\tau-t)} u((1-k_i)W_\tau - B_\tau) I_{\{\tau < \tau'\}} + e^{-\delta(\tau'-t)} u((1-k_i)W_{\tau'} - B_{\tau'}) I_{\{\tau \geq \tau'\}} \}, \quad (27)$$

subject to equations (3), (15), and

$$\frac{dW_s}{W_s} = \mu_i ds + \sigma_i dZ_{i,s}, \quad t \leq s < \min\{\tau, \tau'\}, \quad (28)$$

where τ' is the random future time at which a liquidity shock arrives. This differs from the decision problem in (16) in that it imposes the market clearing condition (26): 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 (27) in the following proposition. The proof is in the Appendix.

Proposition 3: Unless forced to exit the stock market by a liquidity shock, an investor in the economy described above 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, $g_* > 1$. 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

$$U_i(g_t) = \begin{cases} bg_t^{\gamma_1} + \frac{\rho\lambda(1-k_i)}{\rho+\delta-\mu_i}g_t - \frac{\rho\lambda}{\rho+\delta} & \text{if } g_t \in (0, \frac{1}{1-k_i}) \\ c_1g_t^{\gamma_1} + c_2g_t^{\gamma_2} + \frac{\rho(1-k_i)}{\rho+\delta-\mu_i}g_t - \frac{\rho}{\rho+\delta} & \text{if } g_t \in (\frac{1}{1-k_i}, g_*) \\ (1-k_i)g_t - 1 & \text{if } g_t \in (g_*, \infty) \end{cases}, \quad (29)$$

where γ_1 and γ_2 are given by

$$\gamma_1 = \frac{1}{\sigma_i^2} \left[\sqrt{\left(\mu_i - \frac{1}{2}\sigma_i^2\right)^2 + 2(\rho+\delta)\sigma_i^2} - \left(\mu_i - \frac{1}{2}\sigma_i^2\right) \right] > 0 \quad (30)$$

$$\gamma_2 = -\frac{1}{\sigma_i^2} \left[\sqrt{\left(\mu_i - \frac{1}{2}\sigma_i^2\right)^2 + 2(\rho+\delta)\sigma_i^2} + \left(\mu_i - \frac{1}{2}\sigma_i^2\right) \right] < 0, \quad (31)$$

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

$$c_2 = \frac{(\lambda-1)\rho(1-k_i)^{\gamma_2}(\mu_i\gamma_1 - \rho - \delta)}{(\gamma_1 - \gamma_2)(\rho + \delta - \mu_i)(\rho + \delta)} \quad (32)$$

$$(\gamma_1 - 1)c_1g_*^{\gamma_1} + (\gamma_2 - 1)c_2g_*^{\gamma_2} = \frac{\delta}{\rho + \delta} \quad (33)$$

$$c_1\left(\frac{1}{1-k_i}\right)^{\gamma_1} + c_2\left(\frac{1}{1-k_i}\right)^{\gamma_2} = b\left(\frac{1}{1-k_i}\right)^{\gamma_1} + \frac{(\lambda-1)\mu_i\rho}{(\rho + \delta - \mu_i)(\rho + \delta)} \quad (34)$$

$$c_1g_*^{\gamma_1} + c_2g_*^{\gamma_2} + (1-k_i)g_*\frac{\mu_i - \delta}{\rho + \delta - \mu_i} = -\frac{\delta}{\rho + \delta}. \quad (35)$$

To determine μ_i , the equilibrium expected return of stock i , we require that the value function satisfy the conditions in (26), namely $V_i(W, W) = 0$, or equivalently, $U_i(1) = 0$. The parameter μ_i therefore satisfies

$$b + \frac{\rho\lambda(1-k_i)}{\rho + \delta - \mu_i} - \frac{\rho\lambda}{\rho + \delta} = 0. \quad (36)$$

Since the parameters δ , ρ , and λ are constant across investors, μ_i is constant over time, as assumed earlier. In Section 5.2, we use the model described in this section to illustrate the asset pricing applications of realization utility.

5 Applications

Our model may be helpful for thinking about a wide range 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 pricing (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 (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 fail to capture important aspects of the data.

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 (2008) show that realization utility, with no time discounting but with a prospect theory functional form for utility, can predict a disposition effect. This paper proposes a related but distinct view of the disposition effect, namely that it arises from realization utility coupled 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 (2008) point out one possible extra ingredient: a prospect theory functional form for utility, in other words, a utility function that is concave over gains and convex over losses. Such a functional form indeed explains the expediting of gains and the postponement of losses. 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 perhaps the most obvious, in the sense that it is very clear how the effect follows from our initial assumptions. We emphasize, however, that 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 clever 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 is driven by 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: they have a greater propensity to sell stocks trading at a gain relative to purchase price, rather than stocks trading at a loss. 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, before transaction costs, the average return of the individual investors in their sample is on par with the returns on a range of benchmarks, but that, after transaction costs, it falls *below* the benchmark returns. This last finding 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 experience 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:

$$G(s) = 1 - e^{-\rho s} + e^{-\rho s} \left[N \left(\frac{-\ln g_* + \left(\mu - \frac{\sigma^2}{2} \right) s}{\sigma \sqrt{s}} \right) + e^{\left(\frac{2\mu}{\sigma^2} - 1 \right) \ln g_*} N \left(\frac{-\ln g_* - \left(\mu - \frac{\sigma^2}{2} \right) s}{\sigma \sqrt{s}} \right) \right] \quad (37)$$

The expression within the square parentheses in (37) is the probability that g_t reaches g_* in the interval $(0, s)$. Equation (37) therefore has a simple interpretation. 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_* within $(0, s)$. 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 (19)-(22) 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

$$(\mu, \sigma, \delta, k, \rho, \lambda) = (0.03, 0.5, 0.08, 0.01, 0.1, 1.5).$$

Some of the results in Figure 5 are not 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 catch 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 catch 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 very 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 annual turnover rate is 75% and the median annual turnover rate, 30%. 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 0.5.

Underperformance *before* transaction costs

Some studies find that the average individual investor underperforms benchmarks even

before transaction costs (Odean, 1999). 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 relatively 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 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 much 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 which, 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 could potentially sell during this interval are those

targeting liquidation points below \$45. But the majority of these investors *will have sold the stock previously*, before February 28th, when the stock first reached \$45. As the stock moves above \$45 on March 31st, however, investors targeting liquidation points of \$45 and higher 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.⁸

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 much less significant predictor of the propensity to sell. To see this, consider, as before, a stock which is trading at \$30 on January 1st and then rises 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 are targeting 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.⁹

⁸It 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.

⁹The 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, liquidation 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."

5.2 Asset pricing

Our model may also be helpful for understanding certain asset pricing patterns. We now discuss two such applications.

The negative volatility premium

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.

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 low average returns.

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

$$(\delta, \rho, \lambda) = (0.08, 0.1, 1.5), \tag{38}$$

and assume that the transaction cost is the same for all stocks, namely $k = 0.01$. For values of σ ranging from 0.01 to 0.9, we use the equilibrium condition in (36) to compute the expected return that a stock with any given standard deviation must earn in order for markets to clear.

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.

A counterfactual prediction of the top-left graph is that the aggregate equity premium is negative. One possible way of generating a negative relationship between expected return and volatility in the cross-section in conjunction with a *positive* equity premium is to suppose that investors apply different decision rules to different components of their wealth. Suppose

that investors use a standard concave utility function to allocate most of their wealth between a risk-free asset and a stock market index; but that, for the remainder of their wealth – the “play” money in their brokerage accounts which they allocate across individual stocks – realization utility preferences apply. We conjecture that, in such a model, the aggregate equity premium will be determined by the concave utility function and will therefore be positive, but that cross-sectional pricing will be determined by realization utility, thereby leading to a negative cross-sectional relationship between expected return and volatility. We leave this as an open question, one that we hope to study in future research.

Heavy trading of overvalued 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, pushing up the standard deviation of returns σ . 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 may 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 (38) and assume that the transaction cost is the same for all stocks, namely $k = 0.01$. For values of σ ranging from 0.01 to 0.9, we again use condition (36) to compute the corresponding equilibrium expected return; but this time, as a guide to the intensity of trading, we also compute the probability of trade in (37).

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

5.3 Testable predictions

In Sections 5.1 and 5.2, we argue that realization utility offers a simple way of understanding a range of financial phenomena. In most cases, we did not foresee the link between realization utility and the applications we discuss: we saw the link only after completing our analysis. Nonetheless, 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 trades more frequently when transaction costs are lower, is not unique to our model. Four other predictions, however, are more novel: The investor holds stocks with a higher average return for longer, before selling them. Stocks with higher volatility, however, are sold more quickly. The more impatient the investor is, the more often he trades. And the more sensitive he is to losses, the less he trades.

The prediction relating how long a stock is held to its 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 likely that some individual investors *perceive* them to have high average returns.

The prediction relating how long a stock is held to its volatility is easier to test. Indeed, after making this prediction, we found that the answer is already available in the literature. Zuckerman (2006) reports that the individual investors in the Barber and Odean (2000) database *do* hold more volatile stocks for shorter periods of time before selling them.

Our predictions relating trading frequency to investor impatience and investor sensitivity to losses are harder to test, but by no means impossible to test. The difficulty here is obtaining estimates of impatience and sensitivity to losses. 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 impatience and loss sensitivity on the one hand, and trading frequency 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. 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, since realization utility generates Ang et al.’s (2006) negative volatility premium, this premium should be stronger among stocks that are held and traded by individual investors. And fourth, since realization utility generates a 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. For the purpose of judging realization utility, these results are preliminary at best. They do suggest, however, that a more systematic test of our model may be worthwhile.

6 Conclusion

We study the possibility that, aside from standard sources of utility, investors also derive utility from *realizing* gains and losses on assets that they own. We propose a tractable 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 negative premium to volatility in the cross-section, and the heavy trading of highly valued assets. Underlying some of these applications is one of our model’s more novel predictions: that, even if the form of realization utility is linear or concave, investors can be *risk-seeking*.

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:

$$\begin{aligned} & V(W_t, B_t) \\ = & \max \{ u((1-k)W_t - B_t) + V((1-k)W_t, (1-k)W_t), \\ & (1 - \rho dt) E_t [e^{-\delta dt} V(W_{t+dt}, B_{t+dt})] + \rho dt [u((1-k)W_t - B_t)] \} \end{aligned} \quad (39)$$

$$\begin{aligned} = & \max \{ u((1-k)W_t - B_t) + V((1-k)W_t, (1-k)W_t), \\ & E_t [e^{-\delta dt} V(W_{t+dt}, B_{t+dt})] + \rho dt [u((1-k)W_t - B_t) - V(W_t, B_t)] \}. \end{aligned} \quad (40)$$

The first expression on the right-hand side of (39) 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 second expression on the right-hand side shows what happens if the investor instead holds his position for an infinitesimal period dt . 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 (40), cancelling the B_t factor from both sides, and applying Ito's lemma gives

$$\begin{aligned} U(g_t) = & \max \{ u((1-k)g_t - 1) + (1-k)g_t U(1), \\ & U(g_t) + \left[\frac{1}{2} \sigma^2 g_t^2 U''(g_t) + \mu g_t U'(g_t) - (\rho + \delta) U(g_t) + \rho u((1-k)g_t - 1) \right] dt \} \end{aligned} \quad (41)$$

Equation (41) implies that any solution to (8) must satisfy

$$U(g_t) \geq u((1-k)g_t - 1) + (1-k)g_t U(1) \quad (42)$$

and

$$\frac{1}{2} \sigma^2 g_t^2 U''(g_t) + \mu g_t U'(g_t) - (\rho + \delta) U(g_t) + \rho u((1-k)g_t - 1) \leq 0. \quad (43)$$

Formally speaking, the decision problem in (8) is an optimal stopping problem. To solve it, we first construct a function $U(g_t)$ that satisfies conditions (42) and (43) 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 (8).

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 (41) is maximized by the second term within the curly brackets and condition (43) 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 (41) is maximized by the first term within the curly brackets and condition (42) 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

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

The solution to this equation is

$$U(g_t) = a g_t^{\gamma_1} + \frac{\rho(1 - k)}{\rho + \delta - \mu} g_t - \frac{\rho}{\rho + \delta} \quad \text{for } g_t \in (0, g_*), \quad (44)$$

where γ_1 is given in equation (12) and where a is determined below.

In the **liquidation** region, we have

$$U(g_t) = (1 - k)g_t(1 + U(1)) - 1. \quad (45)$$

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 (45), 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 (44) that

$$U(1) = a + \frac{\rho(1 - k)}{\rho + \delta - \mu} - \frac{\rho}{\rho + \delta}. \quad (46)$$

The value function must be continuous and smooth around the liquidation point g_* . This implies

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

Solving these two equations, we obtain the expression for a in (13) and the following nonlinear equation for g_* :

$$(\gamma_1 - 1) \left(1 - \frac{\rho k(\rho + \delta)}{\delta(\rho + \delta - \mu)} \right) g_*^{\gamma_1} - \frac{\gamma_1}{1 - k} g_*^{\gamma_1 - 1} + 1 = 0. \quad (47)$$

Equation (47) has a unique solution in the range $(1, \infty)$. To see this, define

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

The parameter restriction in (9) implies that $\gamma_1 > 1$ and that $1 > \frac{\rho k(\rho + \delta)}{\delta(\rho + \delta - \mu)}$. 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

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

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 (8) and cancelling the B_t factor reduces the stopping problem to

$$\begin{aligned} U(g_t) = \max_{\tau \geq t} E_t \{ & e^{-\delta(\tau - t)} [u((1 - k)g_\tau - 1) + (1 - k)g_\tau U(1)] I_{\{\tau < \tau'\}} \\ & + e^{-\delta(\tau' - t)} u((1 - k)g_{\tau'} - 1) I_{\{\tau \geq \tau'\}} \}. \end{aligned} \quad (48)$$

We first verify that the function $U(g_t)$ summarized in equation (11) satisfies conditions (42) and (43). 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 (11) is a convex function. It must therefore lie above the straight line $f_1(g)$ for all $g < g_*$. Condition (42) is therefore satisfied.

We now check that condition (43) holds. Define

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

Note that for $g < g_*$, $H(g) = 0$ by construction. For $g \geq g_*$, $U(g) = f_1(g)$, so that

$$H(g) = -(1 - k)[(\rho + \delta - \mu)(1 + U(1)) - \rho]g + \delta.$$

Substituting (46) and (13) into this expression, we obtain

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

The last equality follows by applying equation (14). Using (12), it is straightforward to show that if $\mu < \rho + \delta$, as assumed in parameter restriction (9), then $\rho + \delta - \mu\gamma_1 > 0$. Therefore, $H(g) < 0$ for $g \geq g_*$. We have therefore confirmed that condition (43) 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

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

The bound on $U'(g)$ implies that the first integral is a martingale, while condition (43) implies that the second integral is non-positive. We therefore have

$$U(g_t) \geq E_t \left[e^{-\delta(\iota-t)} U(g_\iota) \right]. \quad (49)$$

Note also, from condition (42), that

$$E_t \left[e^{-\delta(\iota-t)} U(g_\iota) \right] \geq E_t \left\{ e^{-\delta(\iota-t)} [((1-k)g_\iota - 1) + (1-k)g_\iota U(1)] \right\}. \quad (50)$$

Now consider the expression in the expectation operator of (48). If $\tau \leq \tau'$, then $\iota = \tau$ and

$$\begin{aligned}
&e^{-\delta(\tau-t)} [u((1-k)g_\tau - 1) + (1-k)g_\tau U(1)] I_{\{\tau < \tau'\}} + e^{-\delta(\tau'-t)} u((1-k)g_{\tau'} - 1) I_{\{\tau \geq \tau'\}} \\
&= e^{-\delta(\iota-t)} [((1-k)g_\iota - 1) + (1-k)g_\iota U(1)].
\end{aligned} \quad (51)$$

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

$$\begin{aligned}
& e^{-\delta(\tau-t)}[u((1-k)g_\tau - 1) + (1-k)g_\tau U(1)]I_{\{\tau < \tau'\}} + e^{-\delta(\tau'-t)}u((1-k)g_{\tau'} - 1)I_{\{\tau \geq \tau'\}} \\
& \leq e^{-\delta(\tau-t)}[u((1-k)g_\tau - 1) + (1-k)g_\tau U(1)]I_{\{\tau < \tau'\}} \\
& \quad + e^{-\delta(\tau'-t)}[u((1-k)g_{\tau'} - 1) + (1-k)g_{\tau'} U(1)]I_{\{\tau \geq \tau'\}} \\
& = e^{-\delta(\iota-t)}[(1-k)g_\iota - 1 + (1-k)g_\iota U(1)],
\end{aligned} \tag{52}$$

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

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

where the first inequality follows from (51) and (52), the second from (50), and the third from (49). 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 (41) and inequalities (42) and (43). The only difference is that $u(\cdot)$ now has the piecewise-linear form in (15).

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

$$\frac{1}{2}\sigma^2 g_t^2 U''(g_t) + \mu g_t U'(g_t) - (\rho + \delta)U(g_t) + \rho u((1-k)g_t - 1) = 0. \tag{53}$$

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 (17), 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, $(0, \frac{1}{1-k})$ and $(\frac{1}{1-k}, g_*)$.

For $g_t \in (0, \frac{1}{1-k})$, equation (53) becomes

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

The solution to this equation is

$$U(g_t) = bg_t^{\gamma_1} + \frac{\rho\lambda(1-k)}{\rho+\delta-\mu}g_t - \frac{\rho\lambda}{\rho+\delta} \quad \text{for } g_t \in \left(0, \frac{1}{1-k}\right), \quad (54)$$

where γ_1 is defined in equation (12), and where b is determined below.

For $g_t \in \left(\frac{1}{1-k}, g_*\right)$, equation (53) becomes

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

The solution to this equation is

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

where

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

and where c_1 and c_2 are determined below.

The value function must be continuous and 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 (21), 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 (19), namely

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

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

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

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

$$\begin{aligned} c_1 g_*^{\gamma_1} + c_2 g_*^{\gamma_2} + \frac{\rho(1-k)}{\rho+\delta-\mu}g_* &= (1-k)g_*(1 + U(1)) - \frac{\delta}{\rho+\delta} \\ c_1 \gamma_1 g_*^{\gamma_1-1} + c_2 \gamma_2 g_*^{\gamma_2-1} + \frac{\rho(1-k)}{\rho+\delta-\mu} &= (1-k)(1 + U(1)). \end{aligned}$$

Since, from equation (54),

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

we obtain equation (22),

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

and equation (20),

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

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 (27) using a technique very similar to the one employed in the proofs of Propositions 1 and 2. In particular, we replace μ , σ , and k in (41) with μ_i , σ_i , and k_i – the expected return, 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,

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

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:

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

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:

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

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.
- 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. and W. Xiong (2008), "What Drives the Disposition Effect? An Analysis of a Long-standing Preference-based Explanation," *Journal of Finance*, forthcoming.
- 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 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*, forthcoming.

- 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.
- Odean, T. (1998), "Are Investors Reluctant to Realize Their Losses?" *Journal of Finance* 53, 1775-1798.
- Odean, T. (1999), "Do Investors Trade Too Much?" *American Economic Review* 89, 1279-1298.
- 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.
- Zuckerman, R. (2006), "It's not the Trader, It's the Trade: A Cross-sectional Analysis of Individual Traders' Activity," Working paper, Yale University.

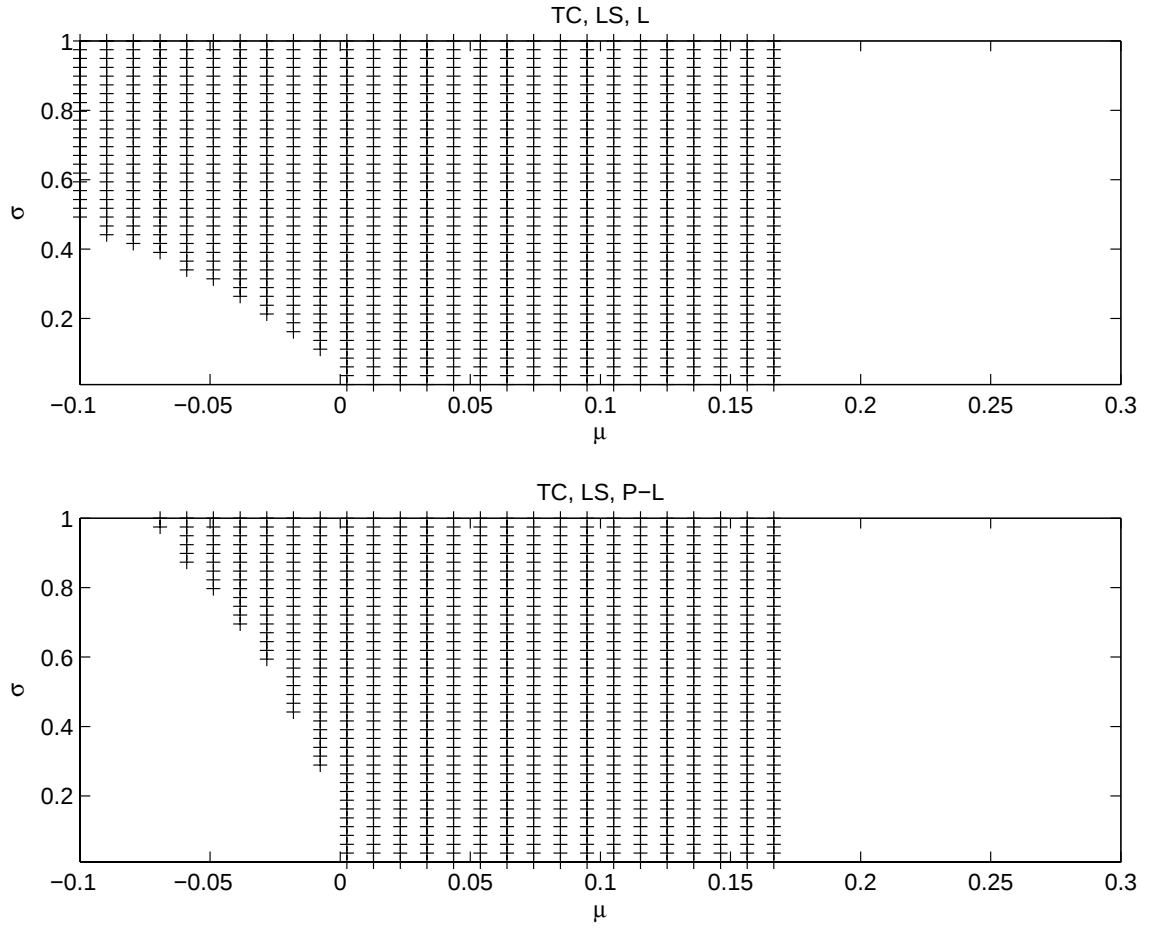


Figure 1. The graphs show, for an investor who derives utility from realized gains and losses, the range of values of a stock's expected return μ and standard deviation σ for which the investor is willing both to buy the stock and to sell it once its price reaches a sufficiently high liquidation point.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

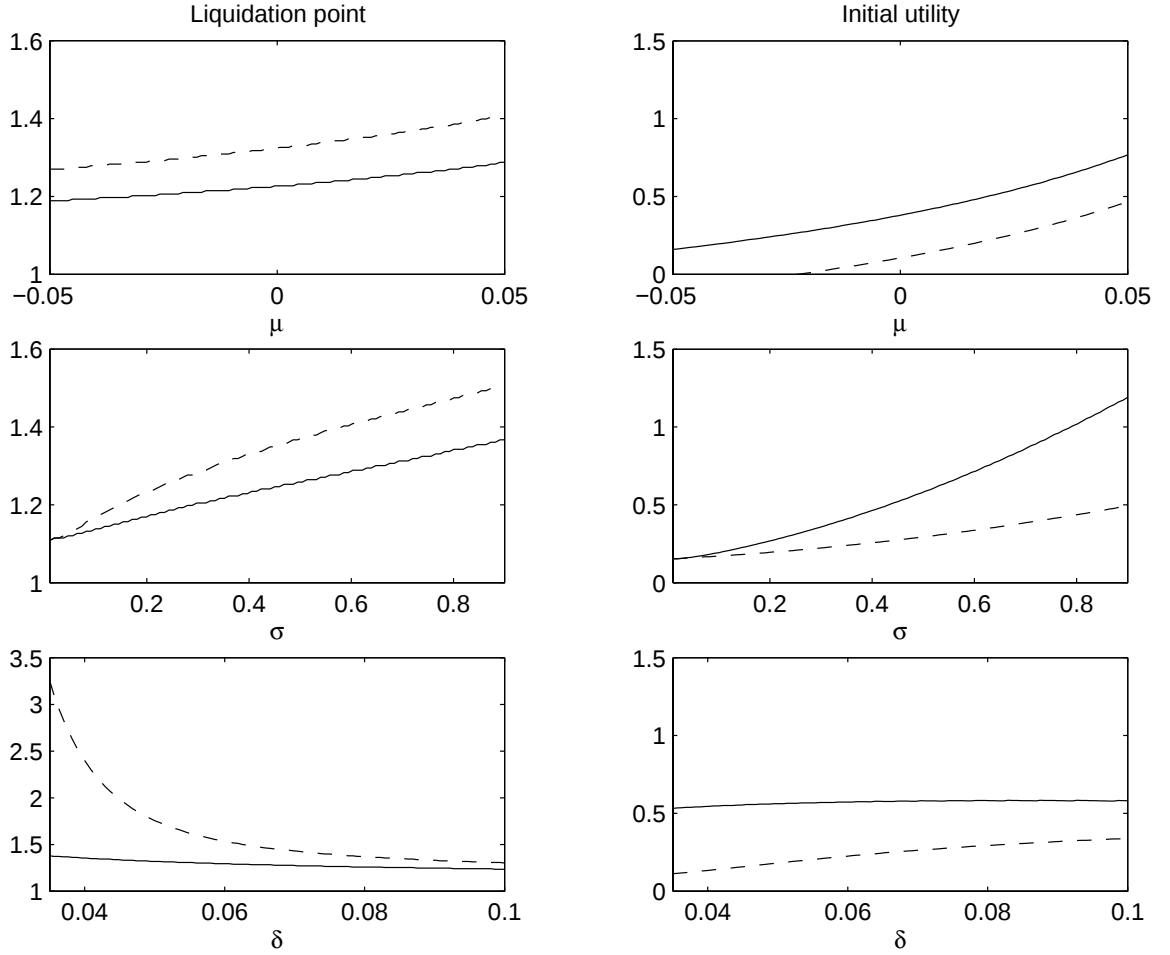


Figure 2. The graphs show, for an investor who derives utility from realized gains and losses, how the liquidation point at which he sells a stock and the initial utility from buying it depend on the stock's expected return μ , its standard deviation σ , and the time discount rate δ . The solid lines correspond to a model that allows for a transaction cost and an exogenous liquidity shock, and in which realization utility has a linear form. The dashed lines correspond to a model that also allows for a transaction cost and an exogenous liquidity shock, but in which realization utility has a piecewise-linear form, so that the investor is 1.5 times as sensitive to realized losses as to realized gains.

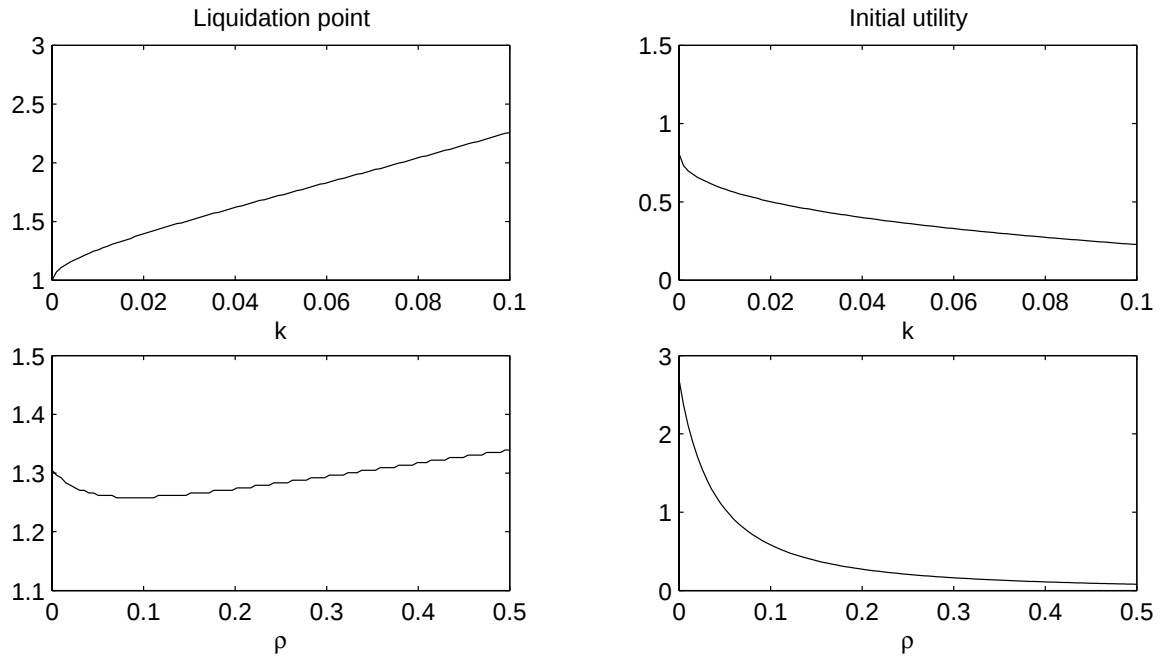


Figure 3. The graphs show, for an investor who derives utility from realized gains and losses, how the liquidation point at which he sells a stock and the initial utility from buying it depend on the transaction cost k and the arrival rate ρ of an exogenous liquidity shock. In these computations, realization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

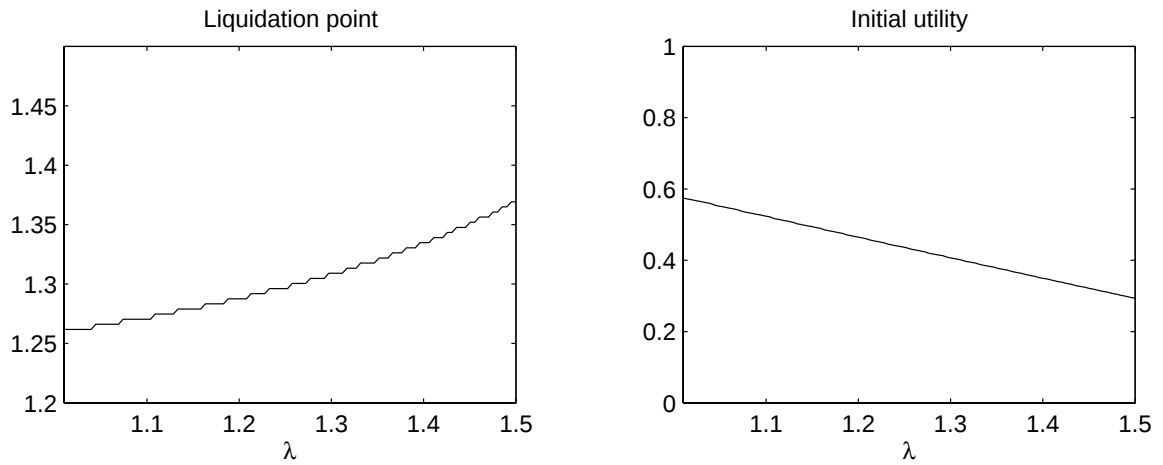


Figure 4. The graphs show, for an investor who derives utility from realized gains and losses, how the liquidation point at which he sells a stock and the initial utility from buying it depend on λ , his relative sensitivity to realized losses as opposed to realized gains. The computations are based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

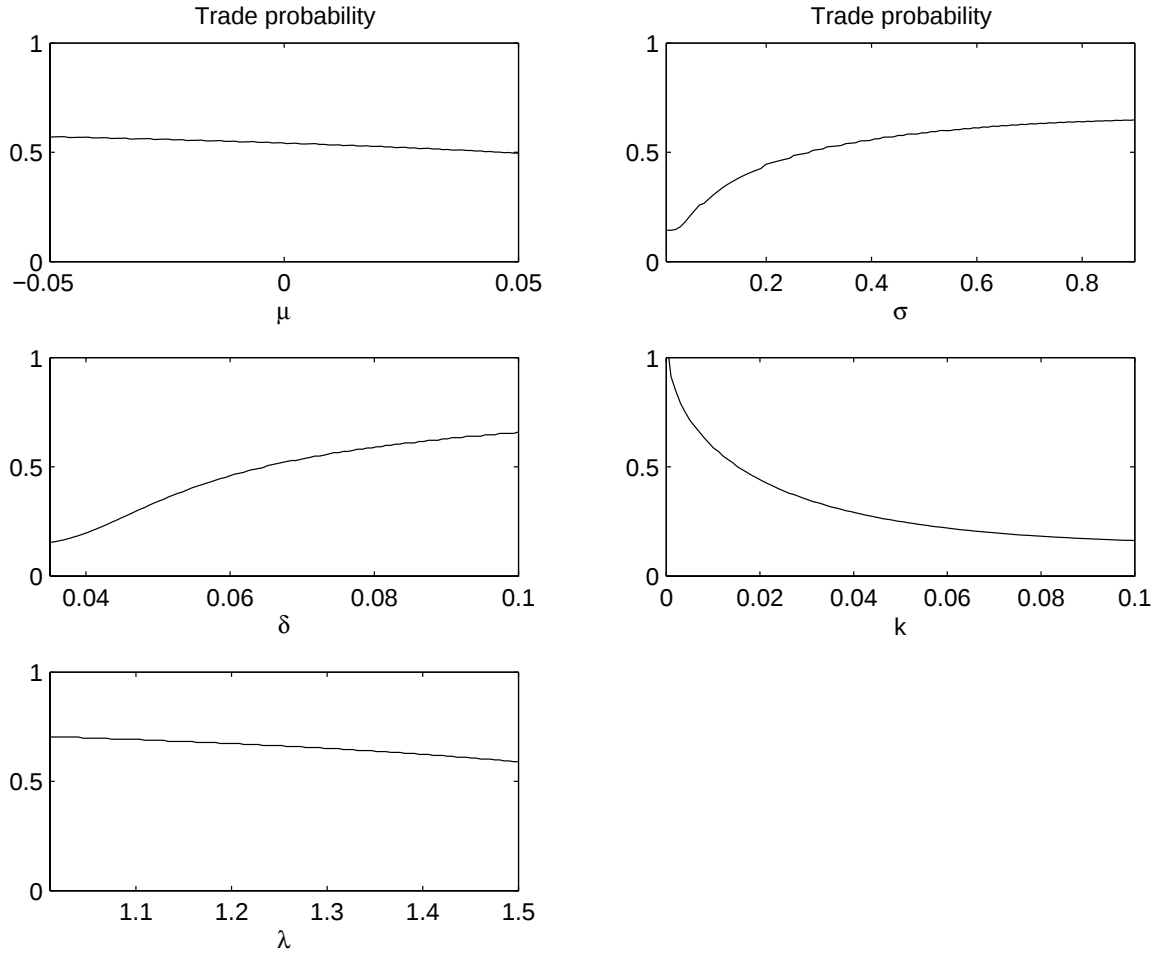


Figure 5. The graphs show, for an investor who derives utility from realized gains and losses, how the probability that the investor will sell a specific stock within a year of buying it depends on the stock's expected return μ , its standard deviation σ , the time discount rate δ , the transaction cost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

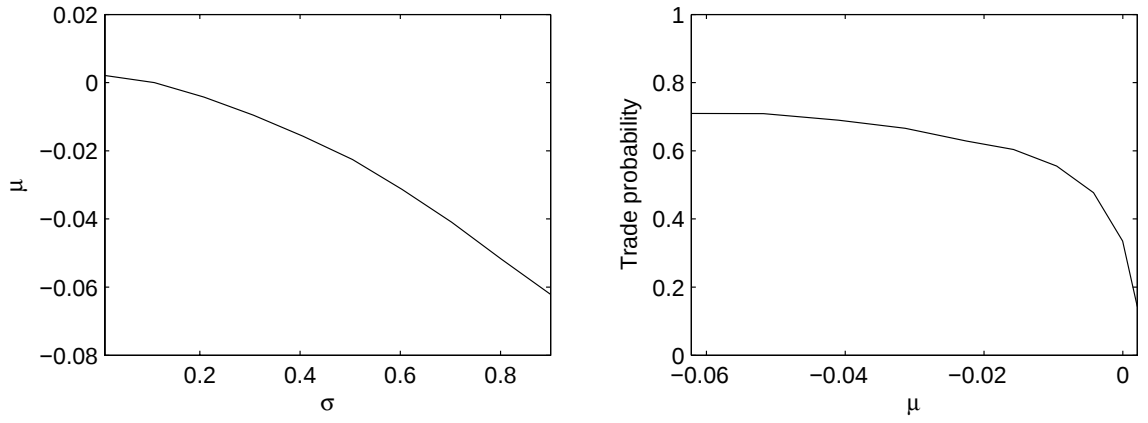


Figure 6. The top-left graph shows, for an economy populated by investors who derive utility from realized gains and losses, the equilibrium relationship between expected return and standard deviation in a cross-section of stocks. The top-right graph shows, for the same cross-section, the equilibrium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