Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Financial Economics 78 (2005) 311–339
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/jfec

# Prospect theory, mental accounting, and momentum☆

Mark Grinblatt[a], Bing Han[b,*]

[a]UCLA Anderson School of Management and NBER, 110 Westwood Plaza, Box 951481, Los Angeles, CA 90095-1481, USA
[b]Department of Finance, The Ohio State University, 2100 Neil Avenue, Columbus, OH 43210-1144, USA

Received 13 December 2002; received in revised form 16 September 2004; accepted 13 October 2004
Available online 2 June 2005

**Abstract**

The tendency of some investors to hold on to their losing stocks, driven by prospect theory and mental accounting, creates a spread between a stock's fundamental value and its equilibrium price, as well as price underreaction to information. Spread convergence, arising from the random evolution of fundamental values and the updating of reference prices, generates predictable equilibrium prices interpretable as possessing momentum. Empirically, a variable proxying for aggregate unrealized capital gains appears to be the key variable that

☆We thank Andrew Lo and Jiang Wang for providing the MiniCRSP database used in this paper, and the UCLA Academic Senate and the Dice Center for Financial Economics at Ohio State University for financial support. We are grateful for the assistance of the editor Bill Schwert and an anonymous referee. We also thank Shlomo Benartzi, Steve Cauley, Bhagwan Chowdhry, Wayne Ferson, Mark Garmaise, Simon Gervais, David Hirshleifer, Harrison Hong, Francis Longstaff, Monika Piazzesi, Richard Roll, and Jiang Wang for invaluable discussions, and seminar participants at Boston College, Emory University, Massachusetts Institute of Technology, the National Bureau of Economic Research, New York University, Ohio State University, Penn State University, University of California at Berkeley, University of California at Irvine, University of California at Los Angeles, University of North Carolina at Chapel-Hill, University of Texas at Austin, University of Washington, University of Wisconsin-Madison, Washington University at St. Louis, the Western Finance Association 2002 Meetings, and the American Finance Association 2003 Meetings for comments on earlier drafts.

*Corresponding author. Tel.: 614 292 1875; fax: 614 292 2418.
*E-mail address:* han.184@cob.osu.edu (B. Han).


doi:10.1016/j.jfineco.2004.10.006

generates the profitability of a momentum strategy. Controlling for this variable, past returns have no predictability for the cross-section of returns.


*JEL classification:* G1; G12

*Keywords:* Prospect theory; Mental accounting; Disposition effect; Profitability of momentum strategies

## 1. Introduction

One of the most well-documented regularities in the financial markets is that investors tend to hold on to their losing stocks too long and sell their winners too soon. Shefrin and Statman (1985) label this the "disposition effect," which has been observed in both experimental markets and financial markets (e.g., stock, futures, options, and real estate), and appears to influence investor behavior in many countries.

Kahneman and Tversky's (1979) theory of choice, "prospect theory," together with Thaler's (1980) "mental accounting" framework, is perhaps the leading explanation for the disposition effect. The main element of prospect theory is an S-shaped value function that is concave (risk averse) in the domain of gains and convex (risk loving) in the domain of losses, both measured relative to a reference point. Mental accounting provides a foundation for the way in which decision makers set reference points for the accounts that determine gains and losses. The main idea is that decision makers tend to segregate different types of gambles into separate accounts, and then apply prospect theory to each account by ignoring possible interactions.

It is fairly easy to see that if the relevant accounts are profits in individual stocks, the prospect theory and mental accounting (PT/MA) combination generates a disposition effect. The reason is that PT/MA investors are risk averse over gambles for some stocks and locally risk loving over gambles for others. The distinction between risk attitudes towards these two classes of stocks is driven entirely by whether the stock has generated a paper capital gain or a paper capital loss. Due to this difference in risk attitudes, investors subject to PT/MA have a greater tendency to sell stocks that have gone up in value since purchase.

To demonstrate this point, consider Fig. 1, which plots the S-shaped value function of a PT/MA investor for outcomes in a particular stock. Let us analyze how this S-shape alters traditional investment behavior. The curve above the inflection point, which is labelled "Reference Point," has the shape of power utility. For true power utility, the fraction of wealth invested in the stock is increasing in the stock's expected return, but is unaffected by the (initial wealth) starting point. How is this demand function shifted by the substitution of a convex utility function to the left of the inflection point? Comparing a starting position at Point D with Point C in Fig. 1, one can infer that demand is increased more at Point C. If we start from Point D, gambles rarely end up in the convex portion of the curve. Indeed, for any given positive mean return, demand increases as the starting position moves left of point D because gambles experience an increasing likelihood of outcomes in the convex

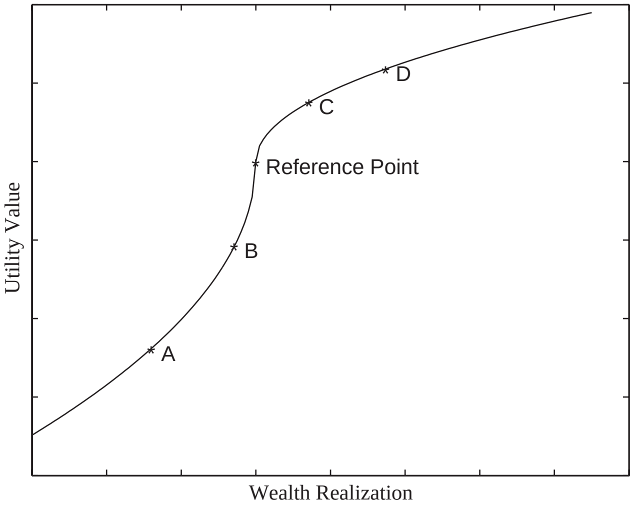


Fig. 1. Prospect theory value function. This figure plots an example of the S-shaped prospect theory value function, generated by

$$U(W) = \frac{(W-R)^{1-\gamma}}{1-\gamma} \quad \text{if } W \geqslant R,$$

$$U(W) = -\lambda \frac{(R-W)^{1-\gamma}}{1-\gamma} \quad \text{if } W < R,$$

where $R$ is a reference level, $\gamma = 0.5$, and $\lambda = 2.25$.

portion of the value function. This pattern of greater demand (for a given mean) as the starting position moves left continues as our starting position crosses the inflection point and moves into the convex region. Clearly, the critical determinant of demand is the starting position in the value function.

When the relevant mental accounts employ the cost basis in a stock as the reference point, the starting positions are dictated by the unrealized capital gain or loss in the stock. Stocks that are extreme winners start the investor at Point D, stocks that are extreme losers start the investor at Point A, and so forth. It follows that a PT/MA demand function differs from that of a standard utility investor not just because winners are less desirable than losers, other things equal. One also concludes that there is a greater appetite for large losers (Point A) than for small losers (Point B). Moreover, there is a lesser desire to shun small winners (Point C) than large winners (Point D) because of the greater degree to which realizations in the convex region enter the expected value calculation.

While the analysis above demonstrates that PT/MA generates the disposition effect, our focus is on the implications of this deviation from rational behavior on asset pricing. This paper considers a model of equilibrium prices in which a group of investors is subject to PT/MA behavior. These investors have demand distortions that are inversely related to the unrealized profit they have experienced on a stock.

Their demand functions distort equilibrium prices relative to those predicted by standard utility theory. The price distortion depends on the degree to which the marginal investor experiences the stock as a winner or a loser. A stock that has been privy to prior good news has excess selling pressure relative to a stock that has been privy to adverse information. If demand for a stock by rational investors is not perfectly elastic, then such a demand perturbation induced by PT/MA tends to generate price underreaction to public information. This produces a spread between the fundamental value of the stock—its equilibrium price in the absence of PT/MA investors—and the market price of the stock. In equilibrium, past winners tend to be undervalued and past losers tend to be overvalued.

The model's price distortions translate into return distortions. To obtain forecastibility in the cross-section of risk-adjusted stock returns, there needs to be a mechanism for undervaluation or overvaluation to diminish over time. Investor heterogeneity is the mechanism the model uses to achieve this. (There are other, more artificial mechanisms that can generate a tendency towards a rational model's valuation over time. A liquidation at a finite horizon is one such alternative mechanism, but we doubt that the effects from such an approach are quantitatively detectable. Dividend streams, a partial liquidation, are subject to the same criticism.) Investor heterogeneity with respect to PT/MA behavior leads to differing demand functions and hence trades of a type consistent with the disposition effect. As disposition effect trading occurs, the cost bases across investors change as does an appropriate aggregate cost basis for investors as a whole. On average, the dynamics of this process tend to reduce the absolute spread between the aggregate cost basis and the market price. Once this reduction in spread occurs, the market price in the next trading round reverts towards its fundamental value. Thus, the PT/MA framework predicts that stocks with paper capital gains will have higher average returns going forward than stocks with paper capital losses.

One implication is that we expect to see momentum in stock returns: any variable that captures the unrealized capital gain experienced by the marginal PT/MA investor will also be a predictor of the cross-section of expected returns. Stocks with high past returns tend to have positive unrealized capital gains for most investors while low past return stocks are more likely to have generated unrealized capital losses.

The model distinguishes itself from others that explain momentum in predicting that (one-period) lagged capital gains are sufficient statistics for forecasting the cross-section of returns. Any other metric of a winner or loser effect will be a noisy proxy for the true capital gain metric. For example, momentum (as well as the disposition effect) can be generated simply by a belief that stock prices revert to a particular value, like the stock price observed one year ago. In such an alternative model, demand pushes the equilibrium price of one-year winners downward, relative to fundamentals, etc. Here, mean reversion is inferred solely from the one-year past return, without reference to the capital gains or losses of investors in each stock. If such an alternative were true, a capital gain-based variable will not be the best predictor of the cross-section of stock returns. Instead, a variable that represents the gap between the current price and the reversion price would dominate as a forecasting variable.

It is the pattern of past returns, combined with the pattern of past trading volume, that determines whether the stock has experienced an aggregate unrealized capital gain or loss. Hence, aggregate capital gains differ from past returns. In our model, proxies for aggregate capital gains (losses) that properly infer aggregate gains from this pattern should be better than past returns as predictors of future returns. We test our model and the importance of PT/MA by running horse races between capital gains and past return variables as predictors of future stock returns.

The empirical implications of our model outlined above are verified with cross-sectional Fama–MacBeth regressions. Motivated by mental accounting, an estimate of the aggregate cost basis for a given stock is used as a proxy for its aggregate reference price. In all of our regression specifications, the capital gains variable thus defined predicts future returns, even after controlling for the effect of past returns, but the reverse is rarely true. Indeed, the return-based momentum effect disappears once the PT/MA disposition effect is controlled for with a regressor that proxies for the aggregate capital gain.

The rest of this paper is organized as follows. In Section 2, we discuss a model that captures the intuition discussed above and we explore its testable implications. Section 3 presents empirical data and provides numerous tests illustrating that our findings are not due to variables used by others in the literature to analyze momentum. Our main finding here is that the capital gains overhang is a critical variable in any study of the relation between past returns and future returns, as the theory predicts. Section 3 also discusses additional implications of the model that have been tested by others. Section 4 concludes the paper.

## 2. The model

This section analyzes how PT/MA-inspired demand functions alter the equilibrium price path of a single risky stock (in an economy with many assets). We assume:

- The risky stock is in fixed supply, normalized to one unit.
- Public news about the date-$t$ fundamental value of the stock, $F_t$, arrives just prior to the date-$t$ round of trading. The fundamental value is the fully rational price that would prevail if there were no PT/MA behavior in the economy.
- The fundamental value follows a random walk, that is

  $$F_{t+1} = F_t + \varepsilon_{t+1}. \tag{1}$$

  This equation generates a convenient benchmark for analyzing the PT/MA-induced alteration of the fully rational price path. With appropriate mental accounts for drift, or if the drift is paid out as a dividend, any other benchmark for fully rational price dynamics would generate identical findings about the price path alteration induced by PT/MA behavior.

The economy has two investor types, where one is not subject to the PT/MA demand distortion. This construction is a simple way of representing the investor heterogeneity needed for reference price updating. It also has the virtue of demonstrating that rational investors cannot undo the equilibrium. The PT/MA investors, a fixed fraction $\mu$ of all investors, have relatively greater (lesser) demand for stocks on which they have experienced losses (gains). The assumed demand functions are

$$\text{Rational demand}: \ \mathrm{D}_t^{\text{rational}} = 1 + b_t(F_t - P_t), \tag{2}$$

$$\text{PT/MA demand}: \ \mathrm{D}_t^{\text{PT/MA}} = 1 + b_t[(F_t - P_t) + \lambda(R_t - P_t)], \tag{3}$$

where $P_t$ is the price of the stock; $R_t$, known prior to date-$t$ trading, is a reference price relative to which PT/MA investors measure their gains or losses; $\lambda$ is a positive constant that measures the relative importance of the capital gain component of demand for PT/MA investors; and, the function $b_t$ represents the slope of the rational component of the demand functions for the stock.

To obtain closed-form solutions for the equilibrium, the PT/MA investor type exhibits a constant geometric perturbation of the rational type's demand function. This modelling device allows us to avoid solving for the rational demand function. Instead, we obtain a closed-form solution for the deviation of a stock's market price from the equilibrium price that would prevail if everyone were rational. This is fully appropriate if we only wish to study the marginal effect of PT/MA behavior on the time-series properties of any equilibrium price path. The process by which the market arrives at a fundamental value in an intertemporal multiasset economy can be quite complicated, but that is not our concern.

In this regard, it is useful to think of $b_t$ as being whatever solves for the optimal rational demand function given a utility function. Eq. (2) does not generally imply linear demand because $b_t$ can be a complex function, depending, for example, on how the return properties of all investments affect utility. The solution to rational investor demand may affect the fundamental value; beyond this, however, it is not relevant to the model. The irrelevance of $b_t$ to all but the fundamental value allows one to alternatively define $b_t$ as the solution to the equilibrium demand of rational investors that have full knowledge of the existence of PT/MA disposition investors. An example in which we explicitly solve for such $b_t$ in a multiperiod exponential utility model for a single-asset market is available from the authors. The existence of an equilibrium here illustrates that arbitrageurs do not fully counter the effect of PT/MA behavior on the equilibrium, even when they are aware of it.

Consistent with the limits to arbitrage argument, we assume $b_t$ is finite. The assumption that rational agents' demands are not perfectly elastic is consistent with every utility function and every numerical simulation we explored in unpublished research undertaken in connection with this study. This assumption generally arises from the risk aversion in utility functions, but it also may reflect liquidity, incomplete information, capital constraints, or other forces restraining unlimited trade by investors. See Shleifer and Vishny (1997) for a thorough discussion of this issue. Among others, Harris and Gurel (1986), Shleifer (1986), Loderer et al. (1991), Kaul et al. (2000), and Wurgler and Zhuravskaya (2002) all provide empirical support for finite price elasticity.

By aggregating investors' demand functions and clearing the market, we find that the equilibrium market price is a weighted average of the fundamental value and the reference price:

$$P_t = wF_t + (1-w)R_t \quad \text{where } w = \frac{1}{1+\mu\lambda}. \tag{4}$$

Since $0<w<1$, the market price underreacts to public information about the fundamental value, holding the reference price constant. The degree of under-reaction, measured by the deviation of $w$ from 1, depends on the proportion of PT/MA investors, $\mu$, and the relative intensity of the demand perturbation induced by PT/MA, $\lambda$. The fewer the number of PT/MA investors, and the smaller the degree to which each perturbs demand, the closer the market price will be to its fundamental value.

Each PT/MA investor is assumed to use a mental account that is separate for each stock. If the relevant reference price is the cost basis for the shares he acquired of that stock, that reference price is updated as shares are exchanged between the investor-types each period. New reference prices are thus weighted averages of old reference prices and the prices at which new shares trade:

$$R_{t+1} = V_t P_t + (1-V_t)R_t. \tag{5}$$

This means that the reference price has a tendency to revert to the current market price. We believe that the updating weight, $V_t$, should be related to the stock's turnover ratio, since the cost basis is the reference price that motivates the mental account. However, our theoretical results would generalize if another mechanism for reference price updating were equally plausible.

With $w$ a constant, the dynamics of the market price can be expressed as

$$P_{t+1} - P_t = w(F_{t+1} - F_t) + (1-w)(R_{t+1} - R_t). \tag{6}$$

Expected changes in $F$ are zero (by definition), while Eq. (5) implies that expected changes in $R$ are of the same sign as the gain—the difference between the market price and the reference price. In the absence of a mechanism for the reference price to change, such as that in Eq. (5), there is no expected price change. However, heterogeneity in the degree to which investors are subject to PT/MA of any variety induces trades and revises the cost basis of the shares in an investor's portfolio.[1] This process of trading redefines the unrealized gains and losses of investors who trade in the stock. When we aggregate across investors, we find that news tends to make the market's effective reference price for a stock's aggregate capital gain converge to the stock's market price. Moreover, since the market price is a weighted average of the fundamental value and reference price (Eq. (4)), the reference price updating also

[1] A contemporaneous theoretical paper by Weber and Zuchel (2001) argues that a single-asset market with a representative investor possessing demand that is linear in mean/variance as well as the deviation of a fixed reference price from the market price will exhibit positive return autocorrelation. With a finite horizon, information about the final liquidation payoff gets more precise over time. The assumed impact of the PT/MA behavior thus decreases monotonically, and the stock price converges deterministically to the fundamental value.

leads both the market price and the reference price to revert to the fundamental value.

Eq. (6) suggests that the expected change in the stock's price from $t$ to $t+1$ is proportional to the change in the reference price that has been generated by trading at date $t$. This, in turn, depends on the size of the unrealized capital gain and the fraction of shares that just changed hands. That is, from Eqs. (5) and (6),

$$\mathrm{E}_t[P_{t+1} - P_t] = (1-w)V_t(P_t - R_t), \tag{7}$$

which is equivalent to

$$\mathrm{E}_t\left[\frac{P_{t+1} - P_t}{P_t}\right] = (1-w)V_t\,\frac{P_t - R_t}{P_t}. \tag{8}$$

This equation suggests that a stock's expected return is monotonically increasing in the marginal investor's (percentage) unrealized capital gain, $(P_t - R_t)/P_t$. Also, for a fixed-sized gain or loss, high current turnover implies that the forecasted absolute return is larger. This is because with high current turnover, next period's unrealized gain or loss is likely to be smaller, shifting next period's aggregate demand function closer to the rational benchmark. This reduction in the PT/MA demand distortion generates an end-of-period equilibrium price that is closer to the fundamental value, giving rise to a larger forecasted absolute return.

Eq. (8) also has implications for momentum in stock returns. Since a stock's capital gain is likely to be correlated with its past return, the past return is a noisy proxy for the unrealized aggregate capital gain that PT/MA investors experience in a stock. With reasonable parameters, our model can generate the empirically observed momentum profit. An earlier draft of this paper demonstrates this, and also contains a nontrivial analytic proof that momentum in stock returns will arise in our model. The proof uses the law of iterated expectations and recursively applies Eqs. (5) and (6).

The model also suggests that the portfolio formation horizon over which momentum is likely to be strongest is an intermediate one. We have confirmed the hump shape of the intensity of the momentum effect as a function of past return horizon with numerical simulations of the model. However, the intuition for the horizon effect is very simple. If the portfolio formation horizon is very short, extreme decile portfolios constructed from stocks with extreme returns can only have small differences in their capital gains and losses. The flow of information over short horizons is often too small to generate large differences in capital gains (or returns) across stocks. The top- and bottom-decile past return performers have larger differences in past returns the longer the past return horizon. However, the spreads for capital gains within these same extreme return decile portfolios do not exhibit the same monotonicity with respect to horizon length. The tendency for the gain, $P_t - R_t$, to revert to zero is quite strong at long horizons: Positions in large losers are replenished with additional shares at more recent market prices and winners tend to be sold. Hence, because trading updates reference prices, there is very little dispersion in the paper gain or loss in the top- and bottom-decile past return performers over a long past return horizon.

## 3. Empirical analysis

We test the theoretical model's price dynamics, expressed in Eq. (8), by analyzing the relation between aggregate capital gains and the cross-section of expected returns. The model also suggests multiplying the gain by one-period lagged turnover. The observed empirical relation between this product and the cross-section of returns is essentially the same as the results with the gain alone as the return predictor. We largely opt for the more parsimonious representation using the gain alone as the key regressor, although we report results with this product variable as the critical return predictor later in this section.

Lacking information on who the PT/MA investors are, we simply estimate a proxy for the market's cost basis in a stock and assume it is the relevant reference price for the mental account. Our estimate of this critical variable is

$$R_t = \sum_{n=1}^{\infty} \left( V_{t-n} \prod_{\tau=1}^{n-1} [1 - V_{t-n+\tau}] \right) P_{t-n}, \tag{9}$$

where $V_t$ is date $t$'s turnover ratio in the stock. The term in parentheses multiplying $P_{t-n}$ is a weight, and all the weights sum to one. The weight on $P_{t-n}$ is the probability that a share was last purchased at date $t-n$ and has not been traded since then. Note that we obtain the same equation by iteratively applying Eq. (5). The cost basis for the market used in empirical work is thus consistent with the reference price dynamics expressed in the model.

Our empirical work utilizes weekly returns, turnover (weekly share trading volume divided by the number of outstanding shares), and market capitalization data from the MiniCRSP database. The data set includes all ordinary common shares traded on the NYSE and AMEX exchanges. NASDAQ firms are not available. The sample period, from July 1962 to December 1996, consists of 1,799 weeks, which is the extent of the weekly data sample. Our choice of weekly data arises from the need to have a reasonable proxy for the capital gains overhang in the market. This requires higher frequency data than monthly data provide and transaction prices that are less influenced by market microstructure than daily data provide. Moreover, the volume numbers on the weekly MiniCRSP data set have been revised to make them more reliable. (see, for example, Lim et al., 2003).

### 3.1. Regression description

We analyze the average slope coefficients of weekly cross-sectional regressions and their time-series $t$-statistics, as in Fama and MacBeth (1973). The week-$t$ return of stock $j$, $r_t^j = (P_t^j - P_{t-1}^j)/P_{t-1}^j$, is the dependent variable. Denote by $r_{t-t_2:t-t_1}^j$ stock $j$'s cumulative return from weeks $t-t_2$ to $t-t_1$. The prior cumulative returns over short, intermediate, and long horizons are used as control regressors for the return effects described in Jegadeesh (1990), Jegadeesh and Titman (1993), and De Bondt and Thaler (1985). Regressor $s_{t-1}^j$, the logarithm of firm $j$'s market capitalization at the end of week $t-1$, controls for the return premium effect of firm size. We also control for the possible effects of volume, such as those described in Lee and Swaminathan (2000) and Gervais et al. (2001), by including $\bar{V}^j_{t-52:t-1}$, stock $j$'s average weekly turnover over the 52 weeks prior to week $t$ as a regressor (and in later regressions, interaction terms, computed as the product of the former volume variable and extreme quintile return rank dummies). We then study the coefficient on $g^j_{t-1}$, a capital gains-related proxy. Formally, we analyze the regression

$$r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 s + a_6 g, \tag{10}$$

and variants of it, where for brevity we drop $j$ superscripts and $t$ subscripts.

Our proxy for the capital gains overhang at the beginning of week $t$ is

$$g_{t-1} = \frac{P_{t-2} - R_{t-1}}{P_{t-2}}. \tag{11}$$

Theory says that this key regressor should employ $P_{t-1}$ instead of $P_{t-2}$. We lag the market price by one week to avoid confounding market microstructure effects, such as bid-ask bounce.

We estimate the aggregate reference price, $R_t$, based on Eq. (9). Obviously, it is not practical to use an infinite sum. Recognizing that distant market prices have little influence on the reference price, we truncate the estimation at five years and rescale the weights to sum to one. This allows us to estimate the reference price in a consistent manner across the sample period. The five-year cutoff, while arbitrary, admits a reasonable portion of our sample period: July 1967 forward. If, for any particular week, a stock lacks at least five years of prior return and turnover data, that stock is excluded from the cross-sectional regression for that week. We verified that our regression results remain about the same when return and turnover data over three or seven prior years are used to calculate the aggregate reference price.

### 3.2. Summary statistics

Fig. 2 plots the weekly time series of the 10th, 50th, and 90th percentile of the cross-section of the capital gains overhang of stocks traded on the NYSE and AMEX. It indicates that there is wide cross-sectional dispersion in this regressor and a fair amount of time-series variation as well. The time-series mean (median) of the difference between the 90th percentile and 10th percentile of the cross-section of the capital gains variable between July 1967 and December 1996 is 76% (60%). For most firms, the time series of this variable exhibits significant comovements with the past returns of the S&P 500 index. The correlations of the above three capital gains percentiles with the past one-year percentage change in the S&P 500 index are, respectively, 0.50, 0.60, and 0.62.

Table 1 Panel A reports summary statistics on each of the variables used in the regression described above. These include time-series means and standard deviations of the cross-sectional averages of the dependent and independent variables, along with time-series means of their 10th, 50th, and 90th percentiles. We obtain further insight into what determines the critical capital gains variable by regressing it (cross-sectionally) on stock $j$'s cumulative return and average weekly turnover for three past

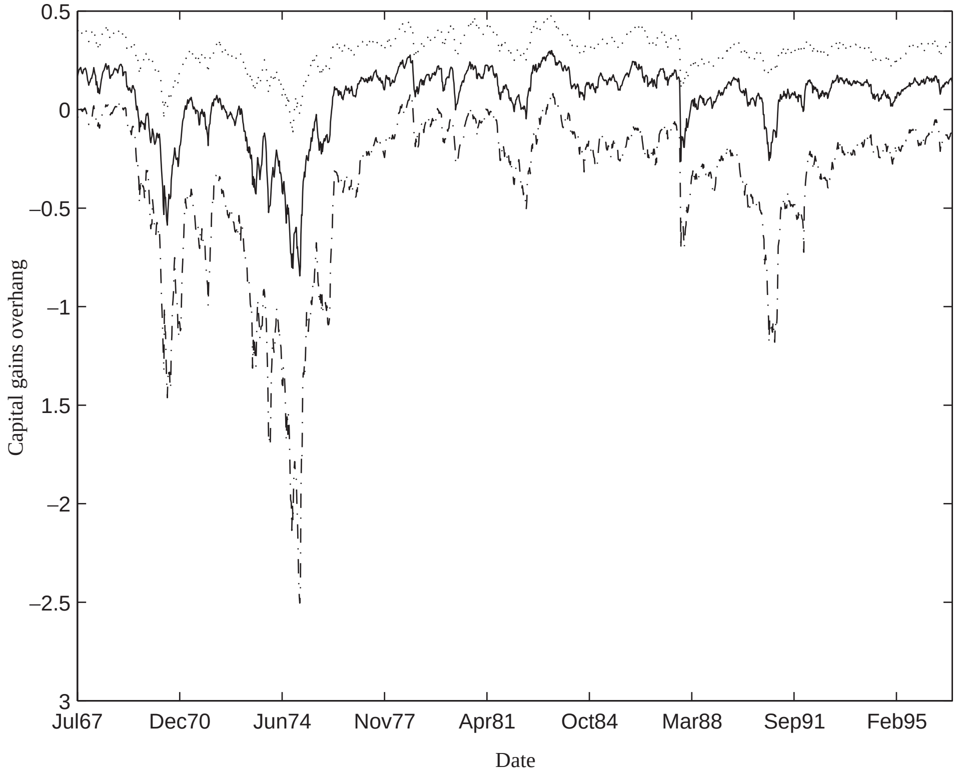


Fig. 2. Time series of cross-sectional percentiles of the capital gains regressor. This figure plots the time series of the empirical 10th, 50th, and 90th percentiles of the cross-sectional distribution of the capital gains regressor. The sample period is from July 1967 to December 1996, for a total of 1,539 weeks. Each week, we include all stocks listed on the NYSE and AMEX that have at least five years of historical trading data from mini-CRSP (used to compute the capital gains variable $g$), where $g$ is one less the ratio of the beginning of week $t-1$ reference price to the end of week $t-2$ price. The week $t-1$ reference price is the average cost basis obtained from the formula

$$R_{t-1} = \frac{1}{k}\sum_{n=1}^{260}\left(V_{t-1-n}\prod_{\tau=1}^{n-1}[1 - V_{t-1-n+\tau}]\right)P_{t-1-n}$$

with $k$ a constant that makes the weights on past prices sum to one.

periods: Short horizon (defined as the last 4 weeks), intermediate horizon (between 5 and 52 weeks ago), and long horizon (between 53 and 156 weeks ago). Size also is included as a regressor.

Panel B of Table 1 reports that, on average, about 59% of the cross-sectional variation in the capital gains variable can be explained by differences in past returns, past turnover, and firm size. What accounts for this explanatory power? Section 2 noted that the reference price is always trying to catch up to a market price that deviates from the reference price for large return realizations. Moreover, the higher the turnover, the faster the reference price converges to the market price. Consistent

Table 1
Summary statistics

This table reports summary statistics of weekly data on NYSE and AMEX securities from July 1967 to December 1996, obtained from the mini-CRSP database. Panel A provides time-series averages of the cross-sectional mean, median, standard deviation, as well as the 10th, 50th, and 90th percentiles for each variable used in the regression

$$r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + +a_5 s + a_6 g,$$

where $r$ is the week-$t$ return; $r_{-t_1:-t_2}$ is the cumulative return from week $t - t_1$ through $t - t_2$; $\bar{V}$ is the average weekly turnover ratio (share volume divided by the number of outstanding shares) over the prior 52 weeks; $s$ is the natural logarithm of market capitalization measured (in thousands of dollars) at the beginning of week $t$; and $g$ is the capital gains regressor, computed as one less the ratio of the beginning of week $t - 1$ reference price to the end of week $t - 2$ price, where the week $t - 1$ reference price is the average cost basis calculated from the formula

$$R_{t-1} = \frac{1}{k} \sum_{n=1}^{260} \left( V_{t-1-n} \prod_{\tau=1}^{n-1} [1 - V_{t-1-n+\tau}] \right) P_{t-1-n},$$

with $k$ a constant that makes the weights on past prices sum to one. Panel B presents more detailed data on the association between the capital gains regressor and other variables. It contains the time-series average of the coefficients and their associated time-series $t$-statistics for 1,539 weekly Fama–MacBeth type cross-sectional regressions of the form

$$g = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 V_{-4:-1} + a_5 V_{-52:-5} + a_6 V_{-156:-53} + a_7 s,$$

where $V_{-t_1:-t_2}$ is the average weekly turnover from $t - t_1$ through $t - t_2$. $R^2_{\text{adj}}$ is the average of the weekly cross-sectional regression $R^2$s adjusted for degrees of freedom.

*Panel A*: *time series average of summary statistics of the regressors*
$r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 s + a_6 g$

| | $r_{-4:-1}$ | $r_{-52:-5}$ | $r_{-156:-53}$ | $\bar{V}$ | $s$ | $g$ |
|---|---|---|---|---|---|---|
| Mean | 0.0119 | 0.1493 | 0.3487 | 0.0092 | 18.7207 | 0.0560 |
| Median | 0.0045 | 0.0940 | 0.2098 | 0.0072 | 18.7251 | 0.1062 |
| Std. dev. | 0.1073 | 0.4192 | 0.7585 | 0.0079 | 1.9441 | 0.2508 |
| 10th percentile | −0.0959 | −0.2538 | −0.3227 | 0.0025 | 16.1399 | −0.2810 |
| 90th percentile | 0.1223 | 0.5816 | 1.1097 | 0.0181 | 21.2322 | 0.3122 |

*Panel B*: *average coefficients and t-statistics* (*in parentheses*) *for the regression*
$g = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 V_{-4:-1} + a_5 V_{-52:-5} + a_6 V_{-156:-53} + a_7 s$

| $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ | $R^2_{\text{adj}}$ |
|---|---|---|---|---|---|---|---|
| 0.5527 | 0.4907 | 0.1771 | −0.9159 | −6.4051 | −2.7843 | 0.0504 | 0.5879 |
| (73.0290) | (51.7965) | (37.5209) | (−7.6351) | (−45.0322) | (−27.8215) | (55.9642) | |

with these facts, Panel B shows that our capital gains variable is positively related to past returns and is negatively related to past turnover. (The time-series mean, median, and standard deviation of the cross-sectional correlation between a stock's capital gains overhang and past one-year return are 0.5482, 0.5529, and 0.1250, respectively.) Controlling for past returns, a low volume winner has a larger capital gain. Also, consistent with our explanation of why intermediate horizons are most important, we find that the effect of intermediate horizon turnover on the capital gains variable is much stronger than the effect of turnover from the other two horizons. Finally, the size coefficient in this regression is significantly positive, perhaps reflecting large firms having grown in the past at horizons not captured by our past return variables and thus tending to have experienced larger capital gains.

### 3.3. Double sorts

Recall that in our model, the risk-adjusted expected return of a stock is determined only by its capital gains overhang. Past returns, which are correlated with the capital gains variable, also predict risk-adjusted returns, but should be noisier predictors. As an initial test of this implication, we study the average returns of portfolios obtained by double sorting both on past one-year returns and the capital gains overhang variable. The double sort is done in two ways. In Panel B of Table 2, stocks are first sorted by their past one-year return into five portfolios labelled as $R1$ (losers), ..., $R5$ (winners). Within each past return quintile, stocks are further sorted into five portfolios by their capital gains overhang from the lowest to the highest quintile $G1, \ldots, G5$. Panel C reverses the sort order.

Table 2 Panel A reports the time-series average of the cutoff values for the capital gains quintiles within each past one-year return sort, and the cutoff values for the past one-year return quintiles within each capital gain sort. The capital gain and past 1-year return are positively correlated, but there is substantial independent variation.

Panels B and C of Table 2 report the average returns of 25 equally weighted portfolios formed on the two double sorts. Januarys are reported separately from non-January months. Consistent with our model's prediction, Panel B shows that during non-January months, for each given past return quintile, the average returns of portfolios increase monotonically with their capital gains overhang quintile. Moreover, the differences between the returns of the highest and lowest capital gains quintiles within each of the past return quintiles is generally significant, ranging from about 0.12% to 0.25% per week (about 6–13% per annum). Panel C indicates that the reverse is not true: The difference between extreme winner and loser quintile portfolios within a given capital gains quintile is generally not significant.

We classify a week as belonging to a particular calendar month if it ends in that month. If we exclude the 30 weeks during our sample that begin in January and end in February from the calculation of average portfolio return during February to December, the lone insignificant $t$-statistic in the right half of Panel B's bottom row also becomes significant.

The portfolio returns during the January months are not consistent with a stable PT/MA parameter $\lambda$. Within Panel B's past return quintiles, the January returns of high capital gains stocks tend to be below those of the low capital gains stocks. This may reflect a December tax-loss selling effect, as we discuss later. It also may reflect a size effect, since the capital gains variable loads positively on the size of the firm. Double sorting cannot explicitly control for other variables that influence the expected return and it is impractical to sort on three or more variables.

To control for these alternative hypotheses, we employ regressions to further test our model, analyzing December and January separately from February through November.

Table 2
Portfolios double sorted on past returns and capital gains

At the beginning of each week $t$, all stocks traded on NYSE and AMEX with five years of prior data are double sorted in two ways. In one double sort, stocks are first sorted into quintiles ($R1 =$ losers, $R5 =$ winners) based on the cumulative return from week $t-52$ through $t-1$. Then within each past return quintile, stocks are further sorted into five equally weighted portfolios by their capital gains $g$ ($G1 =$ lowest, $G5 =$ highest), where $g$ is computed as one less the ratio of the beginning of week $t-1$ reference price to the end of week $t-2$ price. The week $t-1$ reference price is the average cost basis calculated from the formula

$$R_{t-1} = \frac{1}{k} \sum_{n=1}^{260} \left( V_{t-1-n} \prod_{\tau=1}^{n-1} [1 - V_{t-1-n+\tau}] \right) P_{t-1-n},$$

with $k$ a constant that makes the weights on past prices sum to one. The left half of Panel A reports the average capital gains for the 25 portfolios thus obtained, and Panel B reports the average weekly return for these portfolios separately during the January and non-January months. The second double sort reverses the sort order, and the corresponding results are reported in right half of Panel A and Panel C. The sample period is from July 1967 to December 1996; $t$-statistics are reported in parentheses.

*Panel A: time series average of gain/past return for cutoff-percentiles of double sorts*

| Percentile | Gain for cutoff percentile | | | | | Past 1-year return for cutoff percentile | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | R1 | R2 | R3 | R4 | R5 | G1 | G2 | G3 | G4 | G5 |
| 20 | −0.6352 | −0.2553 | −0.0907 | 0.0132 | 0.0948 | −0.3082 | −0.1414 | −0.0273 | 0.0804 | 0.2298 |
| 40 | −0.4011 | −0.1023 | 0.0334 | 0.1216 | 0.2020 | −0.2038 | −0.0516 | 0.0630 | 0.1765 | 0.3638 |
| 60 | −0.2423 | −0.0056 | 0.1103 | 0.1897 | 0.2753 | −0.1059 | 0.0375 | 0.1553 | 0.2795 | 0.5220 |
| 80 | −0.1014 | 0.0834 | 0.1803 | 0.2556 | 0.3525 | 0.0320 | 0.1718 | 0.3019 | 0.4466 | 0.7859 |

*Panel B: mean portfolio return: first sort on past 1-year return*

| | January | | | | | February through December | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | R1 | R2 | R3 | R4 | R5 | R1 | R2 | R3 | R4 | R5 |
| G1 | 0.0280 | 0.0221 | 0.0200 | 0.0190 | 0.0185 | 0.0007 | 0.0010 | 0.0009 | 0.0017 | 0.0015 |
| | (6.6595) | (7.0545) | (6.6434) | (6.9072) | (6.3635) | (0.9502) | (1.5026) | (1.4741) | (2.5651) | (2.0121) |
| G2 | 0.0203 | 0.0133 | 0.0130 | 0.0112 | 0.0108 | 0.0013 | 0.0015 | 0.0019 | 0.0023 | 0.0028 |
| | (5.8051) | (5.2044) | (6.0074) | (4.9437) | (4.4362) | (1.8056) | (2.5014) | (3.3143) | (3.8856) | (3.8803) |
| G3 | 0.0158 | 0.0110 | 0.0097 | 0.0091 | 0.0088 | 0.0010 | 0.0021 | 0.0023 | 0.0026 | 0.0034 |
| | (5.5196) | (4.6749) | (4.8941) | (4.7914) | (3.8287) | (1.3886) | (3.7149) | (4.3071) | (4.7599) | (5.1018) |
| G4 | 0.0133 | 0.0097 | 0.0071 | 0.0058 | 0.0075 | 0.0013 | 0.0020 | 0.0023 | 0.0028 | 0.0036 |
| | (4.9083) | (4.3987) | (3.8552) | (3.2502) | (3.6064) | (1.9823) | (3.7639) | (4.4642) | (5.2575) | (5.4608) |
| G5 | 0.0104 | 0.0065 | 0.0057 | 0.0035 | 0.0062 | 0.0015 | 0.0020 | 0.0026 | 0.0030 | 0.0041 |
| | (4.6832) | (3.5666) | (3.3550) | (2.0306) | (2.9009) | (2.4505) | (3.8347) | (5.1891) | (5.6310) | (6.1472) |
| G5-G1 | −0.0175 | −0.0155 | −0.0143 | −0.0154 | −0.0123 | 0.0008 | 0.0010 | 0.0017 | 0.0012 | 0.0026 |
| | (−6.5141) | (−7.6702) | (−6.2049) | (−7.4544) | (−5.5134) | (1.6146) | (2.6852) | (4.8102) | (3.3838) | (6.7453) |

Table 2 (*continued*)

*Panel C: mean portfolio return: first sort on capital gains*

| | January | | | | | February through December | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | G1 | G2 | G3 | G4 | G5 | G1 | G2 | G3 | G4 | G5 |
| R1 | 0.0225 | 0.0134 | 0.0096 | 0.0064 | 0.0038 | 0.0007 | 0.0015 | 0.0019 | 0.0021 | 0.0026 |
| | (5.5684) | (5.0817) | (4.5359) | (3.5626) | (2.3398) | (0.8731) | (2.2777) | (3.3185) | (4.1417) | (5.2192) |
| R2 | 0.0219 | 0.0130 | 0.0099 | 0.0074 | 0.0036 | 0.0012 | 0.0016 | 0.0022 | 0.0023 | 0.0028 |
| | (6.4120) | (5.0721) | (4.5740) | (4.0176) | (2.0662) | (1.6494) | (2.5961) | (4.1083) | (4.5382) | (5.5472) |
| R3 | 0.0216 | 0.0127 | 0.0107 | 0.0074 | 0.0044 | 0.0012 | 0.0019 | 0.0023 | 0.0025 | 0.0032 |
| | (6.7364) | (5.1222) | (5.2252) | (4.1786) | (2.5006) | (1.7236) | (3.1652) | (4.3337) | (4.8110) | (5.7671) |
| R4 | 0.0221 | 0.0150 | 0.0109 | 0.0086 | 0.0059 | 0.0009 | 0.0017 | 0.0021 | 0.0028 | 0.0033 |
| | (6.7250) | (6.2859) | (4.8773) | (4.4152) | (2.9039) | (1.3786) | (2.6789) | (3.6634) | (4.9085) | (5.1949) |
| R5 | 0.0266 | 0.0166 | 0.0129 | 0.0105 | 0.0089 | 0.0004 | 0.0016 | 0.0025 | 0.0029 | 0.0043 |
| | (7.9147) | (6.3692) | (5.4491) | (4.2918) | (3.5782) | (0.5634) | (2.2953) | (3.5621) | (4.0363) | (5.6412) |
| R5-R1 | 0.0040 | 0.0031 | 0.0032 | 0.0040 | 0.0051 | −0.0003 | 0.0001 | 0.0006 | 0.0008 | 0.0017 |
| | (1.5204) | (1.9304) | (2.2502) | (2.4970) | (2.9409) | (−0.6167) | (0.2956) | (1.4524) | (1.8355) | (3.6596) |

### 3.4. Expected returns, past returns, and the capital gains overhang

Table 3 presents the average coefficients and time-series *t*-statistics for the regression described by Eq. (10) and variations of it that omit certain regressors. Each panel reports average coefficients and test statistics for all months in the sample, for January only, for February through November only, and for December only. (We verify that none of the subsequent results change materially if we exclude 89 ambiguous weeks: (i) Begin in December and end in January; (ii) begin in January and end in February; and, (iii) begin in November and end in December.) All panels include the firm-size regressor. Panel A adds only the three past-return regressors. Panel B adds volume as a regressor to the four regressors from Panel A. Panel C adds the capital gains overhang to the regressors from Panel B.

Panels A and B contain no surprises. As can be seen, when the capital gains overhang variable is excluded from the regression, there is a reversal of returns at both the very short and long horizons, but persistence in returns over the intermediate horizon. Panel B indicates that there is a volume effect, albeit one that is hard to interpret, but it does not seem to alter our conclusions about which horizons offer profitable momentum and contrarian strategies.

Panel C, however, demonstrates that when the capital gains overhang regressor is included in the regression, there is no longer an intermediate horizon momentum effect. The coefficient, $a_2$, is insignificant, both overall and from February through November. However, except for January, there is a remarkably strong cross-sectional relation between the capital gains overhang variable and future returns, with a sign predicted by the model. The estimated average coefficient (0.004) for the capital gains variable from weekly cross-sectional regressions also is consistent with the finding of Jegadeesh and Titman (1993) that momentum strategies generate profits of about 1%/month. Given that the median difference between the 90th and

Table 3
Cross-sectional regression estimates

This table presents the results of Fama–MacBeth (1973) cross-sectional regressions run each week on NYSE and AMEX securities from July 1967 to December 1996. The weekly cross-sectional regressions include all stocks that have at least five years of historical trading data on mini-CRSP. The cross-section of stock returns in week $t$, denoted $r$, are regressed on a constant and some or all of the following variables: $r_{-t_1:-t_2}$ = the cumulative return from week $t-t_1$ through $t-t_2$, computed over three past return horizons; $\bar{V}$ = the average weekly turnover ratio (defined as share volume divided by the number of outstanding shares) over the prior 52 weeks; $s$ =natural logarithm of market capitalization measured (in thousands of dollars) at the beginning of week $t$; and $g$ = the capital gains regressor, computed as one less the ratio of the beginning of week $t-1$ reference price to the end of week $t-2$ price, where the week $t-1$ reference price is the average cost basis calculated from the formula

$$R_{t-1} = \frac{1}{k}\sum_{n=1}^{260}\left(V_{t-1-n}\prod_{\tau=1}^{n-1}[1-V_{t-1-n+\tau}]\right)P_{t-1-n},$$

with $k$ a constant that makes the weights on past prices sum to one. There are a total of 1,539 weekly regressions. The parameter estimates and $t$-statistics (in parentheses) are obtained from the time series of the corresponding cross-sectional regression coefficients. We report the results of regressions over all months, for January only, February through November only, and December only. Panel A omits the capital gains and turnover variables. Panel B omits the capital gains variable. Panel C contains the full set of regressors.

*Panel A*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4\bar{V}$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ |
|---|---|---|---|---|
| All | −0.0482 | 0.0012 | −0.0005 | −0.0004 |
| | (−35.6415) | (2.9527) | (−3.0054) | (−4.2733) |
| Jan | −0.0700 | −0.0087 | −0.0068 | −0.0040 |
| | (−9.6647) | (−4.5972) | (−6.6744) | (−10.9146) |
| Feb–Nov | −0.0459 | 0.0018 | −0.0001 | −0.0001 |
| | (−34.0613) | (4.3344) | (−0.6243) | (−1.4488) |
| Dec | −0.0491 | 0.0051 | 0.0015 | 0.0008 |
| | (−9.9440) | (3.8921) | (2.8930) | (3.0164) |

*Panel B*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4\bar{V} + a_5 s$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ |
|---|---|---|---|---|---|
| All | −0.0488 | 0.0014 | −0.0005 | −0.0540 | −0.0004 |
| | (−37.2470) | (3.5703) | (−2.6700) | (−2.5732) | (−4.4200) |
| Jan | −0.0706 | −0.0086 | −0.0069 | 0.0681 | −0.0042 |
| | (−9.7366) | (−4.5561) | (−6.5561) | (0.9793) | (−11.2309) |
| Feb–Nov | −0.0465 | 0.0021 | −0.0000 | −0.0729 | −0.0001 |
| | (−36.0594) | (5.1324) | (−0.1979) | (−3.1591) | (−1.5202) |
| Dec | −0.0489 | 0.0049 | 0.0015 | 0.0088 | 0.0009 |
| | (−10.2429) | (3.7745) | (2.8046) | (0.1214) | (3.1917) |

Table 3 (*continued*)

| *Panel C*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 s + a_6 g$ | | | | | | |
|---|---|---|---|---|---|---|
| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
| All | −0.0425 | −0.0002 | −0.0007 | −0.0188 | −0.0004 | 0.0040 |
| | (−35.9364) | (−0.6794) | (−5.0871) | (−0.9364) | (−5.2885) | (7.7885) |
| Jan | −0.0520 | −0.0001 | −0.0025 | −0.0620 | −0.0026 | −0.0117 |
| | (−10.9905) | (−0.0477) | (−3.8964) | (−0.9768) | (−8.4381) | (−4.9519) |
| Feb–Nov | −0.0407 | −0.0000 | −0.0006 | −0.0291 | −0.0002 | 0.0050 |
| | (−32.6251) | (−0.0768) | (−3.6950) | (−1.3143) | (−2.8816) | (9.4191) |
| Dec | −0.0498 | −0.0022 | −0.0005 | 0.1238 | 0.0001 | 0.0104 |
| | (−10.8151) | (−1.8953) | (−1.3410) | (1.7980) | (0.2702) | (6.2673) |

10th percentile of capital gains is about 60%, it implies that winners outperform losers by about $0.004 * 60\% = 0.24\%/\text{week}$, or 12.5%/year.

### 3.5. Explaining seasonalities

The seasonalities observed in Table 3 are consistent with the findings of other studies. For example, momentum strategies that form portfolios from past returns over intermediate horizons appear to be most effective in December, and there is a strong reversal in January. (see, for example, Jegadeesh and Titman, 1993; Grundy and Martin, 2001; and Grinblatt and Moskowitz, 2004). Table 3 suggests that these seasonalities are not due to a calendar-based size effect per se. They are fairly easy to explain, however, within the context of our theoretical model if we accept that there is an additional perturbation in demand arising from tax-loss selling.

Odean (1998) and Grinblatt and Keloharju (2001), for example, find that the disposition effect is weakened or even offset in December by the marginal impact of tax-loss selling. A generalized demand function for the PT/MA investor,

$$\mathrm{D}_t^{\mathrm{PT/MA}} = 1 + b_t[(F_t - P_t) + \lambda_t(R_t - P_t)], \tag{12}$$

could plausibly have $\lambda_t$ drift downward in December and revert to its normal positive value in early January. In this case, we would find that the equilibrium effects of this seasonal demand perturbation would be consistent with our empirical findings. The downward drift in $\lambda$ in December implies that market prices move closer to fundamental values. For stocks with capital losses, which imply that the fundamental value is below the market price, convergence towards the fundamental value from the decline in $\lambda$ represents an added force that makes the stock's market price decline even further than it would were $\lambda$ to remain constant. Similarly, the increase in $\lambda$ in early January would make the prices of these same stocks with capital losses deviate again from their fair values, leading to a January reversal.

To understand this more formally, note that with the generalized PT/MA demand, Eq. (12), the expected return, formerly in Eq. (8), generalizes to

$$\mathrm{E}_t\left[\frac{P_{t+1}-P_t}{P_t}\right]=\left[(1-w_t)V_t+\frac{(w_{t+1}-w_t)(1-w_tV_t)}{w_t}\right]\left(\frac{P_t-R_t}{P_t}\right), \tag{13}$$

where $w_t = 1/(1+\mu\lambda_t)$. Hence, if we know that $\lambda_{t+1}$ is going to be lower than $\lambda_t$, which makes $w_{t+1} - w_t$ positive, the expected return between dates $t$ and $t+1$ is going to be larger. The evidence in Grinblatt and Keloharju (2001) suggests that over the course of December, $\lambda$ declines to zero (implying $w_t = 1$), but $\lambda$ is positive during the rest of the year. Viewed from the end of November, this would be like knowing that $w_{t+1}$ is both one and larger than $w_t$, thus generating a larger coefficient on the capital gains regressor in December than would be observed in prior months, during which $w_{t+1} = w_t < 1$. Viewed from the end of December, $w_t$ is one and larger than $w_{t+1}$. This makes the expected price change during January negatively related to the capital gains regressor.

### *3.6. The capital gains variable and volume*

Could the strength of the capital gains variable as a predictor of returns be due to some alternative explanation? Our gain variable is a volume weighting of past returns and many researchers document a connection among volume, past returns, and future returns. Our model's predictions are very specific, however. The largest gain (loss) occurs when there is a lot of volume in the distant past and a large runup (decline) in the stock price with no volume. Because volume is generally quite persistent, it is generally the stocks with low volume that have the most extreme gains for a given past return. If the enhanced precision of the capital gains proxy from the time-series pattern of volume in a stock improves the capital gains variable's forecasting power, that would be evidence in favor of our theory. On the other hand, if the magnitude of the capital gains coefficient in Table 3 Panel C arises entirely from cross-sectional differences in turnover, there could be some alternative explanation for our results. For example, it may be that the most effective trading strategies for momentum involve portfolio formation from past horizons that are more distant for less liquid stocks. This would be picked up by our capital gains variable, but it also would be picked up by a regressor constructed from a reference price that uses average volume and past returns, but ignores the time-series pattern of volume for each stock.

To investigate this issue, we formulate a reference price using the average turnover over the past year in place of each week's actual turnover. In Panels A and B of Table 4, we compute an alternative week-$t$ reference price using $\bar{V}_t^j$, firm $j$'s average weekly turnover from weeks $t-52$ to $t-1$, for all of the 260 $V$s in the five-year approximation of Eq. (9). Panel A replicates Panel C of Table 3, except that in place of the original capital gains variable, we compute an alternative gains variable using the alternative reference price. As Panel A indicates, using a stock's average turnover for the reference price computation instead of the actual weekly turnover generates a significant coefficient on the alternative gains variable. The results are similar to

Table 4
Alternative explanations
This table investigates alternative explanations for the significance of the coefficient on the capital gains regressor. For Panels A and B, $\bar{g}$ is calculated from a reference price using $\bar{V}_t^j$, firm $j$'s average weekly turnover from weeks $t-52$ to $t-1$ in the formula for the gain variable used in week $t$'s cross-sectional regression. Panel A replicates Panel C of Table 3, replacing our original capital gains variable by $\bar{g}$. In Panel B, the relative significance of the two gain variables are compared by including both as regressors. In Panel C and Panel F, we use the product of the gain variable $g$ with last week's turnover as a regressor, rather than the gain variable itself. Panels D and E investigate whether significance is generated by the capital gains variable being correlated with some interaction between past returns and past turnover. Panel D and Panel E add the interaction of average turnover over the past one year and a dummy for the losers (those ranked in the bottom quintile based on past one-year returns) as a regressor, without and with our original capital gains variable, respectively. The parameter estimates and $t$-statistics (in parentheses) are obtained from the time series of cross-sectional regression coefficients. There are a total of 1,539 weekly regressions.

*Panel A*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 s + a_6 \bar{g}$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| All | −0.0419 | −0.0003 | −0.0008 | −0.0160 | −0.0003 | 0.0043 |
| | (−35.3749) | (−0.9434) | (−5.6612) | (−0.8074) | (−4.3955) | (8.0694) |
| Jan | −0.0511 | −0.0004 | −0.0026 | −0.0553 | −0.0030 | −0.0097 |
| | (−10.8551) | (−0.3277) | (−4.0810) | (−0.8509) | (−9.4107) | (−3.9209) |
| Feb–Nov | −0.0403 | −0.0001 | −0.0007 | −0.0266 | −0.0002 | 0.0051 |
| | (−32.0373) | (−0.3395) | (−4.2724) | (−1.2182) | (−1.8236) | (9.2848) |
| Dec | −0.0488 | −0.0019 | −0.0005 | 0.1250 | 0.0003 | 0.0103 |
| | (−10.7502) | (−1.7329) | (−1.3532) | (1.8159) | (1.2605) | (6.0802) |

*Panel B*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 s + a_6 g + a_7 \bar{g}$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ |
|---|---|---|---|---|---|---|---|
| All | −0.0424 | −0.0003 | −0.0008 | −0.0133 | −0.0003 | 0.0028 | 0.0014 |
| | (−34.7482) | (−1.0308) | (−5.1146) | (−0.6459) | (−3.8746) | (2.4381) | (1.2590) |
| Jan | −0.0524 | −0.0010 | −0.0029 | −0.0193 | −0.0022 | −0.0238 | 0.0154 |
| | (−10.9145) | (−0.7809) | (−4.2183) | (−0.2889) | (−7.7160) | (−4.0727) | (2.5766) |
| Feb–Nov | −0.0405 | −0.0001 | −0.0006 | −0.0260 | −0.0001 | 0.0042 | 0.0007 |
| | (−31.3421) | (−0.2966) | (−3.6772) | (−1.1447) | (−1.8369) | (3.6161) | (0.5650) |
| Dec | −0.0513 | −0.0020 | −0.0004 | 0.1160 | 0.0002 | 0.0152 | −0.0048 |
| | (−10.7503) | (−1.6643) | (−0.9656) | (1.5729) | (0.9717) | (4.5511) | (−1.4017) |

*Panel C*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 s + a_6 V_{-1} * g$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| All | −0.0494 | −0.0000 | −0.0006 | −0.0671 | −0.0004 | 0.4876 |
| | (−40.9516) | (−0.0918) | (−4.1084) | (−3.2748) | (−4.4009) | (15.6377) |
| Jan | −0.0661 | −0.0051 | −0.0046 | −0.0535 | −0.0036 | −0.1685 |
| | (−13.0585) | (−3.6265) | (−6.2913) | (−0.8750) | (−10.7269) | (−1.2841) |
| Feb–Nov | −0.0472 | 0.0004 | −0.0004 | −0.0815 | −0.0002 | 0.5271 |
| | (−36.7643) | (1.0364) | (−2.4415) | (−3.5783) | (−1.7896) | (16.2106) |
| Dec | −0.0545 | 0.0013 | 0.0010 | 0.0582 | 0.0007 | 0.7522 |
| | (−13.3977) | (1.2904) | (2.3554) | (0.8299) | (2.7569) | (6.6208) |

Table 4 (*continued*)

*Panel D*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 \bar{V} * D_{\text{loser}} + a_6 s$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| All | −0.0394 | 0.0013 | −0.0003 | −0.0445 | −0.0447 | −0.0003 |
| | (−32.6763) | (3.8855) | (−1.8161) | (−2.1892) | (−2.9342) | (−3.3180) |
| Jan | −0.0575 | −0.0054 | −0.0054 | 0.0137 | 0.1415 | −0.0034 |
| | (−10.7987) | (−3.4640) | (−5.7927) | (0.2254) | (1.8017) | (−9.9059) |
| Feb–Nov | −0.0374 | 0.0019 | 0.0001 | −0.0586 | −0.0587 | −0.0001 |
| | (−29.8235) | (5.1123) | (0.4711) | (−2.5979) | (−3.7923) | (−0.6463) |
| Dec | −0.0417 | 0.0031 | 0.0010 | 0.0350 | −0.0924 | 0.0007 |
| | (−9.0601) | (2.6911) | (2.2312) | (0.5031) | (−1.7119) | (2.7178) |

*Panel E*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 \bar{V} * D_{\text{loser}} + a_6 s + a_7 g$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ |
|---|---|---|---|---|---|---|---|
| All | −0.0426 | −0.0003 | −0.0007 | −0.0140 | −0.0126 | −0.0004 | 0.0040 |
| | (−35.9361) | (−0.8690) | (−4.9930) | (−0.6927) | (−0.8761) | (−5.3449) | (7.8140) |
| Jan | −0.0521 | −0.0003 | −0.0025 | −0.0575 | −0.0315 | −0.0026 | −0.0119 |
| | (−10.9755) | (−0.2333) | (−3.8984) | (−0.9318) | (−0.4486) | (−8.5155) | (−5.0607) |
| Feb–Nov | −0.0409 | −0.0001 | −0.0005 | −0.0231 | −0.0156 | −0.0002 | 0.0050 |
| | (−32.6142) | (−0.2740) | (−3.5798) | (−1.0314) | (−1.0337) | (−2.9252) | (9.4495) |
| Dec | −0.0500 | −0.0020 | −0.0006 | 0.1171 | 0.0345 | 0.0001 | 0.0106 |
| | (−10.8500) | (−1.7670) | (−1.4337) | (1.7137) | (0.7322) | (0.2477) | (6.4041) |

*Panel F*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 \bar{V} * D_{\text{loser}} + a_6 s + a_7 V_{-1} * g$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ |
|---|---|---|---|---|---|---|---|
| All | −0.0495 | −0.0001 | −0.0006 | −0.0645 | −0.0113 | −0.0004 | 0.4882 |
| | (−40.9412) | (−0.2311) | (−4.0264) | (−3.1323) | (−0.7334) | (−4.4434) | (15.7044) |
| Jan | −0.0664 | −0.0051 | −0.0046 | −0.0533 | 0.0321 | −0.0036 | −0.1567 |
| | (−13.0559) | (−3.6300) | (−6.2585) | (−0.8886) | (0.4736) | (−10.7410) | (−1.1920) |
| Feb–Nov | −0.0473 | 0.0003 | −0.0004 | −0.0786 | −0.0136 | −0.0002 | 0.5276 |
| | (−36.7611) | (0.8679) | (−2.3864) | (−3.4329) | (−0.8206) | (−1.8328) | (16.2763) |
| Dec | −0.0546 | 0.0013 | 0.0010 | 0.0610 | −0.0326 | 0.0007 | 0.7425 |
| | (−13.3992) | (1.3174) | (2.3335) | (0.8625) | (−0.6477) | (2.7592) | (6.5686) |

those of Table 3 Panel C, in that intermediate horizon past returns have no predictive power. Moreover, the coefficients and $t$-statistics on the alternative gains variable are similar to those in Table 3 Panel C.

Table 4 Panel B runs a horse race between the two gains variables. It is identical to Table 4 Panel A, except that our original proxy for stock $j$'s capital gains, as used in Table 3 Panel C, is added as a regressor. The inclusion of this variable eliminates the significance of the alternative gains variable, and its coefficient is about the same size as that in Table 3 Panel C in non-January months. While our original capital gains variable is based on an imperfect model of a stock's actual capital gains overhang in the market, it is probably a more precise estimate of the aggregate capital gain for a stock than the alternative capital gains proxy constructed from average historical turnover. The fact that it "knocks out" the alternative gains variable as a predictor of future returns is consistent with more precise estimates of the aggregate capital gain for a stock being better predictors of its future return.

The literature also documents that complicated interactions between volume and past returns improve forecasting. For example, Lee and Swaminathan (2000) find that high volume losers significantly underperform low volume losers. This result is actually consistent with our model, for which volume is a double-edged sword. High volume in the cross-section tends to reduce capital gains. However, this observation ignores the impact of the time series. Our return prediction, found in Eq. (8), multiplies the gains variable by last period's volume. Hence, the largest absolute predicted return occurs if there is low volume in the distant past and then high volume again just before trading takes place. Through trading, the recent updating of the reference prices of PT/MA investors shifts their demand functions closer to the rational benchmark in the subsequent round of trading. It is this convergence to the rational benchmark that drives stock return predictability. We do not use this variable in our earlier regressions largely out of concern that it could be reinventing the Lee and Swaminathan variable in another form. However, if it were used in Table 3 Panel C in place of the gains variable, it approximately doubles the $t$-statistic, as indicated in Table 4 Panel C. Again, it knocks out the intermediate horizon past return as a predictor of future returns.

Our model's prediction that recent volume as a multiplicative interaction term exacerbates the predictive power of capital gains in the cross-section is consistent with other empirical findings of Lee and Swaminathan (2000). They find that most of the predictive power of variables that interact trading volume with past returns is attributable to recent changes in the level of trading activity. To assess their variable against ours, Panels D, E, and F of Table 4 add a proxy for the critical Lee and Swaminathan variable to the mix of regressors: The product of a dummy variable for being in the lowest quintile of past one-year returns and the average past one-year turnover.

Table 4 Panel D analyzes the impact of the Lee and Swaminathan regressor in the absence of a capital gains regressor. Consistent with Lee and Swaminathan (2000), the volume-loser quintile interaction variable is significantly negative. However, once the capital gains variable is added to the regression, as in Table 4 Panel E, the Lee and Swaminathan variable becomes insignificant, while the capital gains coefficient is still highly significant. In Panel F, our capital gains-volume interaction variable replaces the capital gains variable. Again, the Lee and Swaminathan variable is insignificant.

### 3.7. Robustness checks

To most observers, the first and second half of our sample period present different portraits of the stock market. From July 1967 to March 1982, average returns were low, liquidity was low, and trading costs including commissions were high. The second half of our sample period, April 1982 to December 1996 corresponds to a sea change in the stock market. Beginning in August 1982, average returns and trading volume appeared to explode and trading costs rapidly declined. These subperiods also mark an important turning point in the strength of the firm size effect. In the second half of our sample period, size was far less important as a determinant of return premia. Despite these differences, if our theory is part of the core foundation of equilibrium pricing, there should be little difference in the coefficient on our capital gains regressor. Panels A and B of Table 5, which repeat Eq. (10) for the two subperiods, confirm this hypothesis. There is only about a one standard-error difference between the average coefficients on the capital gains regressor in the two subperiods. In both subperiods, the average coefficient is highly significant and positive, while the average coefficient for the intermediate horizon past return is never significant in the presence of the capital gains variable.[2]

We study numerous alternative variables that might explain our results. For example, the maximum 52-week stock price has also been suggested as a possible reference price (see, e.g., Heath et al., 1999). Table 6 Panel A shows that a capital gains proxy constructed using this reference price in the cross-sectional regressions is significantly positive, and it knocks out intermediate horizon past returns as a predictor of future returns. When we add our original capital gains regressor (calculated using the aggregate cost basis for the reference price, as in Eq. (9)), it turns out that both capital gains variables are significantly positive (see Table 6 Panel B). This is what we would expect if the model is correct and both variables are imperfect proxies for the theoretical variable.

The significant predictive power of capital gains for future returns is not an artifact of the weekly frequency of the cross-sectional regressions. In Table 7 Panel A, the dependent return variable in the Fama–MacBeth cross-sectional regression is the monthly return (in lieu of the weekly return). As can be seen from Panel A of Table 7, which corresponds to the specification in Panel C of Table 3, the capital gains variable is still significantly positively related to next month's return. Moreover, once we control for capital gains, the past intermediate horizon return loses its predictive power.

In all of the regressions discussed so far, the intermediate horizon past return is measured by the return between one year and one month ago. To accommodate the possibility that the past-return effect is more complex, we replace this variable by three distinct past return variables: Between three months and one month ago, between six months and three months ago, and between twelve months and six months ago. Panel B of Table 7 shows that once we control for the capital gains overhang, none of these intermediate past returns variables have significant predictive power for future returns. The seasonal pattern stays the same as in Panel C of Table 3. The same results hold when the intermediate past return regressor is replaced by twelve past returns, each over a four-week period.

[2] Although we do not report this formally in a table, the signs and significance of the capital gains overhang regressor are not drastically altered by restricting the sample to various size quintiles either.

Table 5
Robustness check: subsamples

This table presents the subsample results of Fama–MacBeth (1973) cross-sectional regressions that study the relation between capital gains and expected returns. The cross-sectional regressions are run weekly on NYSE and AMEX securities that have five years of historical trading data on mini-CRSP (used to calculate the aggregate cost basis and capital gains). Panel A reports results of the weekly regressions from July 1967 to the end of March 1982. Panel B corresponds to the sample period from April 1982 to the end of December 1996. The parameter estimates and $t$-statistics (in parentheses) are obtained from the time series of the cross-sectional regression coefficients. We report the results of regressions over all months, for January only, February through November only, and December only.

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| *Panel A: July 1967– March 1982* | | | | | | |
| $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 s + a_6 g$ | | | | | | |
| All | −0.0552 | −0.0005 | −0.0013 | −0.0143 | −0.0007 | 0.0046 |
| | (−31.6943) | (−0.9578) | (−5.7743) | (−0.4054) | (−5.2407) | (6.1793) |
| Jan | −0.0631 | −0.0005 | −0.0045 | −0.1711 | −0.0038 | −0.0123 |
| | (−7.9314) | (−0.2847) | (−4.5862) | (−1.7864) | (−8.4704) | (−4.0505) |
| Feb–Nov | −0.0532 | −0.0004 | −0.0011 | −0.0231 | −0.0004 | 0.0058 |
| | (−29.2124) | (−0.6562) | (−4.2579) | (−0.5866) | (−3.0217) | (7.5394) |
| Dec | −0.0666 | −0.0016 | −0.0007 | 0.2267 | 0.0001 | 0.0102 |
| | (−10.9771) | (−0.8759) | (−1.2674) | (1.9665) | (0.2758) | (4.2340) |

*Panel B: April 1982– December 1996*
$r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 s + a_6 g$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| All | −0.0297 | 0.0000 | −0.0001 | −0.0233 | −0.0002 | 0.0035 |
| | (−20.3628) | (0.1063) | (−0.6985) | (−1.2045) | (−1.8569) | (4.8216) |
| Jan | −0.0401 | 0.0004 | −0.0004 | 0.0540 | −0.0013 | −0.0110 |
| | (−8.9945) | (0.2767) | (−0.4923) | (0.6699) | (−3.6897) | (−3.0077) |
| Feb–Nov | −0.0284 | 0.0003 | −0.0001 | −0.0350 | −0.0001 | 0.0042 |
| | (−18.1436) | (0.6909) | (−0.4204) | (−1.7047) | (−0.8256) | (5.7574) |
| Dec | −0.0325 | −0.0028 | −0.0003 | 0.0177 | 0.0000 | 0.0106 |
| | (−5.1506) | (−1.9620) | (−0.5609) | (0.2447) | (0.0839) | (4.6193) |

### 3.8. Additional implications

Since the earliest drafts of this paper, several papers have produced empirical results that are consistent with our model. For example, our model suggests that expected returns are path dependent. While momentum in stock returns may be an artifact of PT/MA behavior because past returns are correlated with variables such as aggregate capital gains, our model implies that for a given past return, some types of paths will generate higher expected returns than others. Holding past returns constant, the capital gains overhang—the difference between current price and the aggregate cost basis—is larger in magnitude for consistent winners and consistent

Table 6
Robustness check using past one-year high as reference price

This table reports results of Fama–MacBeth (1973) cross-sectional regressions run each week on NYSE and AMEX securities from July 1967 to December 1996, similar to Table 3. The only difference is that here each stock's reference price is taken to be its past 52-week high in computing an alternative capital gains regressor $g^*$. We continue to denote by $g$ the original capital gains overhang computed as $(P_{t-2} - R_{t-1})/P_{t-2}$, where

$$R_{t-1} = \frac{1}{k}\sum_{n=1}^{260}\left(V_{t-1-n}\prod_{\tau=1}^{n-1}[1 - V_{t-1-n+\tau}]\right)P_{t-1-n},$$

with $k$ a constant that makes the weights on past prices sum to one. Panel B compares the two gains variables by including them simultaneously as regressors. There are a total of 1,539 weekly regressions. The parameter estimates and $t$-statistics (in parentheses) are obtained from the time series of cross-sectional regression coefficients. We report the results of regressions over all months, for January only, February through November only, and December only.

*Panel A*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4\bar{V} + a_5 s + a_6 g^*$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| All | −0.0426 | 0.0003 | −0.0003 | −0.0198 | −0.0004 | 0.0043 |
| | (−36.8035) | (0.9040) | (−1.6889) | (−1.0619) | (−4.8157) | (8.9118) |
| Jan | −0.0538 | −0.0034 | −0.0050 | −0.0664 | −0.0031 | −0.0052 |
| | (−11.4558) | (−2.7598) | (−5.9271) | (−1.1459) | (−9.6145) | (−2.4838) |
| Feb–Nov | −0.0411 | 0.0006 | 0.0001 | −0.0257 | −0.0002 | 0.0052 |
| | (−33.5797) | (1.8163) | (0.4991) | (−1.2414) | (−2.4379) | (10.2560) |
| Dec | −0.0454 | 0.0009 | 0.0011 | 0.0834 | 0.0005 | 0.0048 |
| | (−10.3698) | (0.7818) | (2.6272) | (1.3520) | (1.9373) | (3.0944) |

*Panel B*: $r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4\bar{V} + a_5 s + a_6 g + a_7 g^*$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ |
|---|---|---|---|---|---|---|---|
| All | −0.0437 | −0.0005 | −0.0005 | −0.0030 | −0.0004 | 0.0026 | 0.0032 |
| | (−37.6869) | (−1.6377) | (−3.4039) | (−0.1583) | (−5.6575) | (4.8956) | (6.6383) |
| Jan | −0.0528 | −0.0003 | −0.0023 | −0.0690 | −0.0026 | −0.0130 | 0.0028 |
| | (−11.5159) | (−0.2813) | (−3.4364) | (−1.1577) | (−8.6337) | (−5.9260) | (1.5015) |
| Feb–Nov | −0.0422 | −0.0003 | −0.0003 | −0.0085 | −0.0003 | 0.0033 | 0.0037 |
| | (−34.2590) | (−1.0275) | (−1.8000) | (−0.4012) | (−3.3040) | (5.9960) | (7.1028) |
| Dec | −0.0492 | −0.0023 | −0.0008 | 0.1149 | 0.0001 | 0.0114 | −0.0013 |
| | (−11.1069) | (−1.9592) | (−1.8100) | (1.8163) | (0.3849) | (6.4287) | (−0.7723) |

losers. Stocks that are consistent winners or stocks that are at their all-time highs are more likely to have larger unrealized gains than stocks that have the same past return, achieved through a handful of outstanding months in the distant past. Grinblatt and Moskowitz (2004) find that momentum profits are stronger for consistent winners. George and Hwang (2004) find that profits to a portfolio formation strategy based on nearness to a 52-week high are superior to those based on past returns over a fixed horizon.

Table 7
Additional robustness checks
This table provides further robustness checks on the results of Fama–MacBeth (1973) cross-sectional regressions reported in Panel C of Table 3. In Table 3, the dependent variable is the weekly return and the cross-sectional regressions are run each week. Panel A of this table reports the average coefficient estimates and $t$-statistics (in parentheses) obtained from cross-sectional regressions that are run once a month between July 1967 and December 1996, the using monthly return as the dependent variable. Panel B of this table differs from Panel C of Table 3 only in that the past return between one year and one month ago is being replaced by three nonoverlapping returns over intermediate past horizons: months -1 to -3, -4 to -6, and -7 to -12. We report the results of regressions over all months, for January only, February through November only, and December only.

*Panel A*: *monthly regressions*
$r = a_0 + a_1 r_{-4:-1} + a_2 r_{-52:-5} + a_3 r_{-156:-53} + a_4 \bar{V} + a_5 s + a_6 g$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ |
|---|---|---|---|---|---|---|
| All | −0.0674 | 0.0021 | −0.0017 | −0.1708 | −0.0010 | 0.0127 |
| | (−16.6929) | (1.4099) | (−2.5740) | (−2.2182) | (−2.6895) | (5.4241) |
| Jan | −0.0910 | 0.0013 | −0.0089 | −0.3477 | −0.0090 | −0.0345 |
| | (−4.1583) | (0.2239) | (−2.5983) | (−1.4647) | (−5.5341) | (−4.1412) |
| Feb–Nov | −0.0632 | 0.0028 | −0.0011 | −0.1923 | −0.0004 | 0.0152 |
| | (−15.2941) | (1.7591) | (−1.4978) | (−2.2249) | (−1.0585) | (6.3402) |
| Dec | −0.0864 | −0.0043 | −0.0014 | 0.2232 | 0.0006 | 0.0344 |
| | (−7.0285) | (−0.7936) | (−0.9822) | (1.0281) | (0.5383) | (4.6423) |

*Panel B*: *more refined intermediate horizon past returns*
$r = a_0 + a_1 r_{-4:-1} + a_2 r_{-13:-5} + a_3 r_{-26:-14} + a_4 r_{-52:-27} + a_5 r_{-156:-53} + a_6 \bar{V} + a_7 s + a_8 g$

| Period | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ | $a_8$ |
|---|---|---|---|---|---|---|---|---|
| All | −0.0437 | −0.0040 | −0.0004 | 0.0004 | −0.0008 | −0.0211 | −0.0004 | 0.0047 |
| | (−35.9816) | (−5.5374) | (−0.6773) | (1.1588) | (−5.5893) | (−1.1094) | (−5.5158) | (8.8904) |
| Jan | −0.0530 | −0.0027 | 0.0032 | 0.0015 | −0.0024 | −0.0905 | −0.0025 | −0.0123 |
| | (−10.3896) | (−0.8353) | (0.9277) | (1.2120) | (−3.7591) | (−1.5476) | (−8.0866) | (−5.2322) |
| Feb–Nov | −0.0419 | −0.0037 | −0.0005 | 0.0005 | −0.0006 | −0.0272 | −0.0003 | 0.0058 |
| | (−32.9948) | (−4.9480) | (−0.9261) | (1.2087) | (−4.2594) | (−1.2835) | (−3.2591) | (10.6286) |
| Dec | −0.0518 | −0.0083 | −0.0025 | −0.0011 | −0.0006 | 0.1056 | 0.0001 | 0.0112 |
| | (−10.7456) | (−2.7814) | (−1.3900) | (−0.7937) | (−1.5469) | (1.6812) | (0.2205) | (6.5430) |

Our model also makes predictions about trading volume and volatility. Goetzmann and Massa (2003) derive several additional implications of our model for volume and volatility, as well as returns. They find strong empirical support for these implications. For example, in a period of rising prices, there is a significant negative correlation between the prevalence of disposition investor trades and turnover or volatility. Consistent with our model's implication, Goetzmann and Massa (2003) find that a behavioral factor capturing the stochastic change in the percentage of disposition investors is significantly negatively related to returns when the capital gains overhang is positive. Further, their results suggest that exposure to this disposition factor seems to be priced.

In our model, stocks with large unrealized capital gains underreact to positive firm-specific news while stocks with large unrealized capital losses underreact to negative firm-specific news. This should generate post-announcement drift. Frazzini (2004) finds additional support for our model by showing that the post-announcement drift following earnings surprises and changes in analyst recommendations is significantly higher when the news and capital gains overhang have the same sign. Moreover, the magnitude of the post-announcement drift is directly related to the amount of unrealized capital gains/losses experienced by the stockholders prior to the announcement date. Most significantly, Frazzini (2004) finds that a holdings-based proxy for capital gains is a better predictor of returns than both past returns and our turnover-based proxy for capital gains. This supports the prediction of our model that a more precise measure of paper capital gains will better forecast future returns.

## 4. Conclusion

There is a growing literature that shows that the prospect theory and mental accounting frameworks are useful for explaining many asset pricing anomalies. These anomalies range from the equity premium puzzle (Benartzi and Thaler, 1995; Barberis et al., 2001) to the high volatility and long-horizon predictability of stock returns (Barberis et al., 2001) to IPO-related pricing anomalies (Loughran and Ritter, 2002; Ljungqvist and William, 2005; Barberis and Huang, 2004). Motivated by prospect theory and mental accounting, our paper's model of equilibrium asset prices adds to this literature by explaining why momentum exists in the cross-section of stock returns. The model presented here is consistent with the empirical evidence on the disposition effect. It also predicts that the difference between a stock's market price and its aggregate cost basis will be positively related to the stock's expected future return as well as a better predictor of future average returns than past one-year returns.

The empirical tests of our model strongly support its main implications. Using double sorts, we find that holding past returns constant, the average returns of portfolios increase monotonically with their capital gains overhang quintile. On the other hand, there is generally no significant difference between the average returns of portfolios sorted on past returns within each capital gains overhang quintile. Using Fama–MacBeth regressions, we find a significantly positive cross-sectional relation between a stock's capital gains overhang and its future stock return. Moreover, the predictive power of the intermediate horizon past return becomes insignificant when one controls for capital gains. In other words, the Jegadeesh and Titman (1993) momentum effect largely disappears. Our results are robust, and cannot be explained by cross-sectional differences in liquidity or the interaction of past returns and turnover.

In the model developed here, fully rational arbitrageurs cannot eliminate the impact of capital gains on equilibrium prices. Although prices always underreact to news, trying to arbitrage away this underreaction is risky. There are several reasons for this. First, rational investors cannot ascertain when reference prices, and hence market prices, will converge to fundamental values. Market prices can diverge further from their fundamental values before they converge. Second, the fundamental values are unpredictable. Thus, risk-averse rational agents will not take infinite positions to get rid of the mispricing. Third, if rational agents have limited capital or a short horizon, their ability to eliminate the impact of PT/MA behavior on prices will be further reduced. For example, Liu and Longstaff (2004) show that arbitrageurs optimally underinvest or even walk away from an arbitrage opportunity when faced with margin requirements. Moreover, DeLong et al. (1990) show that when there are positive feedback traders in the economy, rational arbitrageurs who anticipate their impact on demand can front-run the positive feedback investors and may even destabilize prices, rather than help to bring market prices in line with fundamental values.

The high risk associated with the strategy of buying stocks with low reference prices and shorting stocks with high reference prices (relative to market prices) applies even when there are many assets. Within a linear factor model, for example, this naive attempt at arbitrage does not account for the fact that the sensitivities to priced and unpriced factors are correlated with discounts/premia related to reference prices. Hence, a portfolio constructed solely on the basis of reference price discounts/premia necessarily has large factor exposure. In empirical work, it would appear as if there is a PT/MA factor.

We mostly focus on the momentum anomaly here. Clearly, there are other behavioral models that seem to address this issue in interesting ways as well. Daniel et al. (1998) present a model in which investors are overconfident and also suffer from a self-attribution bias. Their behavior generates delayed overreaction to information that is eventually reversed. Barberis et al. (1998) argue that the representative heuristic may lead investors to extrapolate current earnings growth well into the future. At the same time, the conservativism bias of investors leads to underreaction to public information. In Hong and Stein (2000), agents can use only part of the information about the economy because of communication frictions. In their model, private information diffuses slowly through the population of investors, which causes underreaction in the short run. Momentum traders can profit by trend-chasing, but cause overreaction at long horizons in doing so. Our model differs from other behavioral models in suggesting that aggregate capital gains is the critical variable in forecasting the cross-section of returns. Our model also differs in its prediction of disposition behavior, a well-documented empirical regularity.[3]

Perhaps the most promising avenue for further research is the implications this model has for volume. Most equilibrium models have no trade in them. Our behavioral model is an exception. We have not explored the volume implications empirically, but they are interesting. Volume turns out to be a path-dependent function of movements in the fundamental value. The greatest volume is found when

[3]See, for example, Shefrin and Statman (1985), Case and Shiller (1988), Ferris et al. (1988), Odean (1998), Weber and Camerer (1998), Heath et al. (1999), Grinblatt and Keloharju (2001), Genesove and Mayer (2001), Shapira and Venezia (2001), Wermers (2003), Kaustia (2004), and Frazzini (2004).

there is a sudden price drop after a large and long-lasting price runup. Although some researchers have begun to study the volume implications of the model, the volume implication discussed above is one of a number of volume implications that deserve investigation.

## References


Barberis, N., Huang, M., 2004. Stock as lotteries: the implication of probability weighting for security prices. Working Paper, University of Chicago.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

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

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

Case, K., Shiller, R., 1988. The behavior of home buyers in boom and post-boom markets. New England Economic Review 1988, 29–46.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53, 1839–1886.

De Bondt, W., Thaler, R., 1985. Does the stock market overreact? Journal of Finance 40, 793–808.

DeLong, J.B., Shleifer, A., Summers, L., Waldmann, R., 1990. Positive feedback investment strategies and destabilizing rational speculation. Journal of Finance 45, 379–395.

Fama, E., MacBeth, J., 1973. Risk, return, and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.

Ferris, S., Haugen, R., Makhija, A., 1988. Predicting contemporary volume with historic volume at differential price levels: evidence supporting the disposition effect. Journal of Finance 43, 677–697.

Frazzini, A., 2004. The disposition effect and under-reaction to news. Yale ICF Working Paper No 4–24.

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

George, T., Hwang, C.Y., 2004. The 52-week high and momentum investing. Journal of Finance 59, 2145–2176.

Gervais, S., Kaniel, R., Mingelgrin, D., 2001. The high-volume return premium. Journal of Finance 56, 877–919.

Goetzmann, W., Massa, M., 2003. Disposition matters: volume, volatility and price impact of a behavioral bias. NBER Working Paper 9499.

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

Grinblatt, M., Moskowitz, T., 2004. Predicting stock price movements from the pattern of past returns: the role of consistency and tax-loss selling. Journal of Financial Economics 71, 541–579.

Grundy, B., Martin, S., 2001. Understanding the nature of risks and source of rewards to momentum investing. Review of Financial Studies 14, 29–78.

Harris, L., Gurel, E., 1986. Price and volume effects associated with changes in the S&P 500: new evidence for the existence of price pressures. Journal of Finance 41, 815–829.

Heath, C., Huddart, S., Lang, M., 1999. Psychological factors and stock option exercise. Quarterly Journal of Economics 114, 601–628.

Hong, H., Stein, J., 2000. A unified theory of underreaction, momentum trading and overreaction in asset markets. Journal of Finance 54, 2143–2184.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881–898.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.

Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263–291.

Kaul, A., Mehrotra, V., Morck, R., 2000. Demand curves for stocks do slope down: new evidence from an index weights adjustment. Journal of Finance 55, 893–912.

Kaustia, M., 2004. Market-wide impact of the disposition effect: evidence from IPO trading volume. Journal of Financial Markets 7, 207–235.

Lee, M.C., Swaminathan, B., 2000. Price momentum and trading volume. Journal of Finance 55, 2017–2069.

Lim, T., Adamek, P., Lo, A., Wang, J., 2003. The trading volume and the MiniCRSP Database: an introduction and user's guide. MIT LFE Working Paper No. LFE-1038–98.

Liu, J., Longstaff, F., 2004. Losing money on arbitrages: optimal dynamic portfolio choice in markets with arbitrage opportunities. Review of Financial Studies 17, 611–641.

Ljungqvist, A., William J., 2005. Does prospect theory explain IPO market behavior? Journal of Finance, forthcoming.

Loderer, C., Cooney, J., Van Drunen, L., 1991. The price elasticity of demand for common stock. Journal of Finance 46, 621–651.

Loughran, T., Ritter, J., 2002. Why don't issuers get upset about leaving money on the table in IPOs? Review of Financial Studies 15, 413–443.

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

Shapira, Z., Venezia, I., 2001. Patterns of behavior of professionally managed and independent investors. Journal of Banking and Finance 25, 1573–1587.

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

Shleifer, A., 1986. Do demand curves for stock slope down? Journal of Finance 41, 579–590.

Shleifer, A., Vishny, R., 1997. The limits to arbitrage. Journal of Finance 52, 35–55.

Thaler, R., 1980. Toward a positive theory of consumer choice. Journal of Economic Behavior and Organization 1, 39–60.

Weber, M., Camerer, C., 1998. The disposition effect in securities trading: an experimental analysis. Journal of Economic Behavior and Organization 33, 167–184.

Weber, M., Zuchel, H., 2001. The Disposition effect and momentum. Working Paper, University of Mannheim.

Wermers, R., 2003. Is money really "smart"? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Working Paper, University of Maryland.

Wurgler, J., Zhuravskaya, E., 2002. Does arbitrage flatten demand curves for stocks? Journal of Business 75, 583–608.