Long-Term Return Reversals: Overreaction or Taxes?

Thomas J. George

tom-george@uh.edu

C.T. Bauer College of Business
University of Houston
Houston, TX 77204

and

Chuan-Yang Hwang

cyhwang@ntu.edu.sg

Division of Banking and Finance
Nanyang Business School
Nanyang Technological University
Singapore 639798

and

Department of Finance
School of Business and Management
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

June 2006

Long-Term Return Reversals: Overreaction or Taxes?

Abstract

Long-term reversals in US stock returns are better explained by the rational reactions of investors to locked-in capital gains than irrational overreaction to news. Predictors of returns based on the overreaction hypothesis have no power, while those that measure the extent of locked-in capital gains do have predictive power and completely subsume past returns measures that traditionally have been used to predict long-term returns. We also examine data from Hong Kong, where investment income is not taxed. Reversals are non-existent in Hong Kong, and returns are not forecastable either by traditional measures, or by measures based on the capital gains lock-in hypothesis that successfully predict returns in US data.

Key words: return reversal, overreaction, capital gains lock-in.

Introduction

Short to medium term momentum in security returns was first documented by Jegadeesh and Titman (1993). They show that winner stocks over the past six months outperform losers by 1% per month during the next six to twelve months. DeBondt and Thaler (1985) show that loser stocks in the past three to five years outperform winners by 25% over the next three years. These findings of short-term momentum and long-term reversals are the empirical cornerstones of the study of behavioral finance. Prominent theoretical models in this area such as Barberis, Shleifer and Vishny (1998), Daniel, Hirshleifer and Subrahmanyam (1998) and Hong and Stein (1999) all treat short-term momentum and long-term reversals as inseparable phenomena. In Barberis, Shleifer and Vishny and Hong and Stein, momentum occurs because traders are slow to revise their priors when new information arrives. Long-term reversals occur because when traders finally do adjust, they overreact. In Daniel, Hirshleifer and Subrahmanyam, momentum occurs because traders overreact to prior information when new information confirms it. Long-term reversals occur as the overreaction is corrected in the long run. In all three models, short-term momentum and long-term reversals are sequential components of the process by which the market absorbs a news item. This view is supported by evidence that return momentum documented in Jegadeesh and Titman (1993) reverses in the long run (see Jegadeesh and Titman (2001)).

Recent studies present evidence that these two phenomena are unrelated. George and Hwang (2004) show that the nearness of a stock's price to its 52-week high dominates the traditional momentum measures such as those used in Jegadeesh and Titman (1993) and

1

Moskowitz and Grinblatt (1999) in predicting returns, and that the momentum captured by their measure does not reverse in the long run. If short-term momentum and long-term reversals are related, then momentum predicted by the dominant momentum measure would reverse. Grinblatt and Han (2005) present a theoretical model in which short-term momentum is a result of the "disposition effect" first documented by Odean (1998): investors tend to continue holding loser stocks while selling winners. In their model, the disposition effect leads to price momentum that does not reverse in the long run.

If short-term momentum and long-term reversals are unrelated, the explanation for why long-term reversals exist becomes an open question. DeBondt and Thaler (1987) argue that their results are consistent with investor overreaction to news. However, they also show that long-term reversals have a very strong seasonal pattern—significant long-term reversals associated with loser stocks occur only in January. This suggests that loser reversals have more to do with tax loss selling than investor overreaction. The strong seasonal nature of reversals is confirmed by Grinblatt and Moskowitz (2004), who also document January reversals for long-term losers.

This paper investigates whether taxes can explain long-term reversals. Since capital gains are taxed only when realized, investors with locked-in gains have an incentive not to sell winners in order to delay paying capital gains taxes. Consequently, investors' reservation prices for the sale of winner stocks are elevated by the benefit of capital gains

deferral. Stocks with large embedded capital gains will have higher prices, and hence lower expected returns, than otherwise identical stocks with no embedded capital gains.

The tax hypothesis we examine is modeled by Klein (1999) in a dynamic general equilibrium framework. In Klein's model, investor heterogeneity generates transactions, but at prices that are higher for stocks where investors have locked-in capital gains than otherwise identical stocks without such gains. Investors' optimal responses to the taxation of capital gains generate long-term reversals. This happens for two reasons. First, as time passes, buyers arrive whose endowments and tastes for risk are such that they are willing to pay the premium prices demanded by those who own stocks with embedded gains. As turnover occurs, the new owners do not have large locked-in gains and are willing to sell without demanding large premiums. Second, as investors with locked-in gains approach the terminal point of their investment horizon, the benefit to further deferring taxes on gains (and hence their reservation selling prices) decreases.

In this paper, we conduct a test of both the overreaction hypothesis and the capital gains lock-in effect to determine which of the two hypotheses better explains long-term reversals using stock returns data from both the US and Hong Kong. Since investment income is not taxed in Hong Kong, using both data sets enables us to isolate the explanation for reversals more clearly than we could with US data alone. In the US data, we find that the predictions of the overreaction hypothesis have no power, while measures of embedded gains do have predictive power and completely subsume the

traditional measures in predicting reversals.   In the Hong Kong data, neither the

overreaction hypothesis nor measures of embedded gains predict reversals.

Our results are consistent with the hypothesis that long-term reversals are not caused by

investor overreaction; rather they reflect a rational response by investors who account for

tax consequences when making their portfolio decisions. Other explanations for long-

term reversals that are consistent with rational behavior, but are unrelated to taxes, are

presented in Berk, Green and Naik (1999), Lewellen and Shanken (2002) and Brav and

Heaton (2002). Furthermore, Conrad, Cooper and Kaul (2003) suggest that evidence of

long-term reversals may be a consequence of data snooping. Our results favor the lock-in

hypothesis over these explanations.

The rest of paper is organized as follows. Section I discusses further the overreaction and

lock-in hypotheses. Section II describes the data and methodology. Results are presented

in section III. Section IV concludes the paper with a brief summary and discussion of our

results in light of rational explanations of long-term reversals that are not based on taxes.

I. Hypotheses

A. Overreaction Hypothesis

The overreaction hypothesis assumes that investors make systematic mistakes when they

react to information. It can arise from biased self-attribution as discussed in Daniel,

Hirshleifer and Subrahmanyam (1998). It can also arise from investors' tendency to form

beliefs about future performance by extrapolating from recent past performance. Upon seeing a stock that has experienced a string of good news or a period of growth, investors may wrongly believe that growth will continue, which pushes the stock's price higher than is justified by the news. The price later reverses when investors realize mistakes were made.

In empirical tests of overreaction, classifying stocks as winners and losers by past return *and* by their nearness to historical highs and lows can have an advantage over classification based on past returns alone. This is because past returns depend on price changes over a fixed historical period (e.g., five years), and a past return might still be large well after the reversal is underway. Consequently, a stock for which very significant good news has arrived can continue to be classified as a winner by past return alone, even after the reversal associated with overreaction has occurred. The price of such a stock would no longer be near its historical high. Alternatively, a stock with a large positive past return whose price is near a historical high is a stock for which very significant news has arrived *and* the reversal associated with overreaction has not yet occurred. If the overreaction hypothesis is true, such a stock should exhibit a strong reversal (negative return) in the future. A symmetric argument applies to stocks with large negative returns whose prices are near historical lows.

This reasoning implies that if overreaction is the explanation for the long-term reversals documented by DeBondt and Thaler (1985, 1987), then among winner (loser) stocks with similar past returns, we should see larger reversals for those whose prices are close to

5

historical highs (lows). Note that the overreaction hypothesis gives symmetric predictions for winners and losers. This is in contrast to the implications of the capital gains lock-in hypothesis discussed below.

B. Capital Gains Lock-In Hypothesis

Our tests of the capital gains lock-in hypothesis are based on the fact that capital gains taxes are paid only when gains are realized. Investors who hold a stock with a large embedded gain will sell only if the price compensates them to forego the value of delaying payment of the tax. Thus, stocks with embedded gains will trade at premiums relative to otherwise identical stocks that do not have embedded gains. As shares eventually turn over to buyers without large embedded gains, the marginal investor's reservation price falls resulting in a slowly dissipating reversal. This intuition is formalized in Klein (1999).

It is important to recognize the asymmetry between the effects of capital gains and losses on investor behavior in empirical work. The empirical tests in Klein (2001), for example, treat the price effects of capital gains and losses as though a capital loss is just a negative capital gain. The lock-in effect of a capital loss will not be opposite to that of a capital gain, however.   There are two reasons for this. First, Constantinides's (1984) tax timing option model predicts that investors should realize capital losses as soon as they occur in order to re-establish short-term status for further losses or to accelerate the starting date to establish long-term status for future gains.   Second, even if investors do not strictly follow the prediction of Constantinides (because of transaction costs or a disposition

6

effect), they need not discount their selling prices of stocks with embedded losses below those of otherwise identical stocks without embedded losses because buyers should stand ready to pay the same price for either stock. Thus, the central prediction of the capital gains lock-in hypothesis is that long-term reversals occur for stocks with embedded capital gains, but no reversals occur for stocks with embedded losses.

II. Data and Methodology

The basic data consist of monthly prices, returns and other characteristics for all NYSE, AMEX and NASDAQ companies covered by CRSP from 1963 through 2001. The Hong Kong data are described later.

As in DeBondt and Thaler (1985, 1987), and Klein (2001), we use a five-year horizon to study long-term reversals. In each month we use information from the past five years to form winner and loser portfolios that are held for the next five years. We study reversals by examining the returns to these portfolios. Significant negative (positive) excess returns to winner (loser) portfolios indicate reversals.

With a 60 month horizon in mind, we denote the price on the last trading day of month t as P_t, and its monthly price history as $P_{t-1}, P_{t-2}, \ldots, P_{t-60}$ (the prices are adjusted for stock splits and stock dividends). We construct the measures below from which to judge whether a given stock belongs in a winner or loser portfolio at time t.

1. **Traditional Return Measure (*TR*):** This measure is simply the stock's return over the portfolio formation period, $\frac{P_t - P_{t-60}}{P_{t-60}}$, plus the return from reinvesting dividends. This is the measure traditionally used in studies of both momentum and long-term reversals. Jegadeesh and Titman (1993) use past six-month or one-year returns to study momentum at short and intermediate horizons. DeBondt and Thaler (1985, 1987) use the past five-year returns to study reversals at long horizons.

2. **Five-Year High Measure (FYH):** This measures the nearness of the month-*t* price to the stock's five-year high:

$$\mathbf{FYH}_t = \frac{1}{\sum_{n=1}^{60} w_{t-n}} \sum_{n=1}^{60} w_{t-n} \left(\frac{P_t - P_{t-n}}{P_{t-n}} \right)$$

where $w_{t-n} = 1$ if $P_{t-n} = \max\{P_t, P_{t-1}, ..., P_{t-60}\}$ and $w_{t-n} = 0$ otherwise. As explained in the previous section, the overreaction hypothesis predicts that both winners and losers reverse. This measure is used to test whether winners reverse. If winner reversals are caused by overreaction, then *TR*-winner stocks whose prices are also near five-year highs (i.e., stocks that are both *TR* and *FYH* winners) should have stronger reversals than *TR* winners whose prices are not near five-year highs. This is because nearness to the five-year high indicates that a reversal due to overreaction has not yet occurred for the former set of stocks.

3. **Five-Year Low Measure (*FYL*):** This measures the nearness of the month-*t* price to the stock's five-year low:

$$\textbf{FYL}_t = \frac{1}{\sum_{n=1}^{60} w_{t-n}} \sum_{n=1}^{60} w_{t-n} \left(\frac{P_t - P_{t-n}}{P_{t-n}} \right)$$

where $w_{t-n} = 1$ if $P_{t-n} = \min\{P_t, P_{t-1}, \ldots, P_{t-60}\}$ and $w_{t-n} = 0$ otherwise. It is used to test whether losers reverse. If loser reversals are caused by overreaction, then *TR*-loser stocks whose prices are also near five-year lows (i.e., *TR* and *FYL* losers) should have stronger reversals than *TR*-losers whose prices are not near five-year lows. As noted in point 6 below, this measure is also used to test the capital gains lock-in hypothesis. However the focus there is on *FYL* winner*s* rather than losers.

The first three measures are designed to test the overreaction hypothesis. The next set of variables is designed to test the lock-in hypothesis. They measure gains embedded in investors' stock holdings. These variables are as consistent with Klein's (1999) model as possible. In Klein's model, the aggregate deferral benefit, which is reflected as a premium in the equilibrium stock price, is the risk-tolerance-weighted average of the benefits of capital gains deferral to individual investors. These individual deferral benefits are increasing and convex in capital gains measured on a per-share basis. Convexity arises because greater capital gains increase the value of deferring taxes for a given horizon investment, but greater gains also cause investors to endogenously increase the expected horizons over which the investment will be held. Assuming investors are homogeneous, and ignoring the convexity, the aggregate deferral benefit reduces to the equally weighted average of gains embedded in investors' stock holdings.

4. **Equal-Weighted Gain and Loss Measure (*EWGL*)**: This measures the average embedded percentage capital gains *net of capital losses,* under the assumption that investors acquire shares uniformly over the past 60 months:

$$\mathbf{EWGL}_t = \frac{1}{60} \sum_{n=1}^{60} \left(\frac{P_t - P_{t-n}}{P_{t-n}} \right)$$

This measure treats gains and losses symmetrically, whereas the capital gains lock-in hypothesis predicts that reversals are associated only with gains. Therefore, if the capital gains lock-in hypothesis is correct, this measure should be dominated by a measure of gains alone in explaining reversals.

5. **Equal-Weighted Gain Only Measure (*EWGO*)**:

$$\mathbf{EWGO}_t = \frac{1}{\sum_{n=1}^{60} w_{t-n}} \sum_{n=1}^{60} w_{t-n} \left(\frac{P_t - P_{t-n}}{P_{t-n}} \right)$$

where $w_{t-n} = 1$ if $P_t > P_{t-n}$ and $w_{t-n} = 0$ otherwise. This measure specifically recognizes the asymmetry between capital gains and losses—i.e., that capital losses do not generate negative deferral benefits. This is still an imperfect measure of embedded gains because investors' share acquisitions are not necessarily uniform across time. Despite the approximation, if the lock-in hypothesis is the true explanation for reversals, then winners based on this measure will predict winner reversals well but losers by this measure should have no power in predicting loser reversals. Furthermore, *EWGO* should dominate other measures such as *TR* and *EWGL* in predicting winner reversals.

A variation of *EWGO* is the volume-weighted gain only measure defined as

$$\text{VWGO}_t = \frac{1}{\sum_{n=1}^{60} w_{t-n}} \sum_{n=1}^{60} w_{t-n} \left(\frac{P_t - P_{t-n}}{P_{t-n}} \right) \quad \text{where } w_{t-n} = volume_{t-n} \text{, if } P_t > P_{t-n} \text{ and}$$

$w_{t-n} = 0$ otherwise. This approximation of the aggregate embedded gain is

equivalent to assuming the number of investors who purchase stocks at P_{t-n} is

proportional to $volume_{t-n}$. However, we find that *VWGO* is much less effective

than *EWGO* in predicting winner reversals, whether it is used alone or in

conjunction with other variables (the tables are omitted for brevity). The

variable $volume_{t-n}$ is trading volume for the entire month. Using this in

conjunction with the month-end price may be too noisy an estimate of the

volume-weighted transaction prices for the entire month to improve on *EWGO*. [1]

6. **Five-Year Low Measure (*FYL*)**: The five-year low measure (*FYL*) defined in

item 3 above can also be used to test the capital gains lock-in hypothesis. While

EWGO measures gains under the assumption that the number of investors

acquiring shares is uniform over time, *FYL* can be thought of as a measure of

embedded gains under the extreme assumption that shares are acquired all at

once at the five-year low. In this respect, *FYL* measures the maximum potential

capital gain. The capital gains lock-in hypothesis therefore predicts that stocks

identified as *winners* based on a ranking by *FYL* will experience negative future

excess returns. This obviously overestimates the magnitude of the embedded gain,

but we include this measure in our tests because *EWGO* ignores the fact that the

deferral benefit is a convex function of embedded capital gains. The maximum

gain alone probably overestimates the aggregate deferral benefit, whereas *EWGO*

probably underestimates it. If, despite their imperfections, *EWGO* and *FYL*

winners (and not losers) dominate *TR, FYH* and *EWGL* in explaining reversals,

we will regard it as evidence supportive of the lock-in hypothesis.

The *TR* measure is coded as missing if either P_t or P_{t-60} is missing. For the other

measures, we require at least 40 non-missing prices in the series $P_{t-1}, P_{t-2},…, P_{t-60}$.

Otherwise, those measures are coded as missing, and the stock is excluded from

consideration for that month.

For each of the measures described above, a stock is deemed a winner (loser) in month t

if its value of the measure ranks in the top (bottom) *Q%* of measures for all stocks in

month t. If, by a given measure, a stock is a winner (loser) in month t, then a winner

(loser) dummy variable for that measure is set to one for that stock in month t. Otherwise,

the dummy variable takes the value of zero. Thus, for every month, each stock has

associated with it a value for winner and loser dummy variables defined with respect to

each of the measures above. Adding a *W* or *L* suffix denotes the dummy variables. For

example, the winner (loser) dummy variable for *TR* is *TRW* (*TRL*). In Table II, $Q = 10\%$

cutoffs are used for consistency with the earlier literature. As described below, less

restrictive cutoffs of $Q = 30\%$ are used for the remaining tests.

We follow the Fama-MacBeth (1973) style regression approach taken by George and

Hwang (2004) and Grinblatt and Moskowitz (2004) to measure and compare returns to

different investment strategies. This approach has the advantage of being able to isolate

the return to a particular strategy while hedging (zeroing out) the impact of other strategies and other variables known to affect returns. It also enables us to decompose a strategy's long-term return into components attributable to sub-periods of the investment horizon.

If an investor forms portfolios of winners or losers every month and holds these portfolios for the next T months, the return earned in a given month t is the equal-weighted average of the returns to T portfolios, each formed in one of the past T months t-j (for $j=1$ to $j=T$). The contribution of the portfolio formed in month t-j to the month-t return can be obtained by estimating a cross-sectional regression of the form:

$$R_{it} = b_{0jt} + b_{1jt} R_{i,t-1} + b_{2jt} size_{i,t-1} + b_{3jt} 52wkhW_{i,t-j} + b_{4jt} 52wkhL_{i,t-j} + b_{5jt} M1W_{i,t-j}$$

$$+ b_{6jt} M1L_{i,t-j} + b_{7jt} M2W_{i,t-j} + b_{8jt} M2L_{i,t-j} + e_{ijt} \qquad (1)$$

where R_{it} is the return to stock i in month t, and $M1W_{i,t-j}$ ($M1L_{i,t-j}$) equals one if stock i is a winner (loser) in month t-j according to measure 1. Dummies $M2W_{i,t-j}$ and $M2L_{i,t-j}$ are defined similarly for measure 2. For example, if measure 1 is five-year low (*FYL*), then the winner and loser dummies are defined based on the lowest price found over the 60 month period ending at t-j—i.e., month ends at $\{t$-j-$60,...,t$-$j\}$. We also include equity market capitalization, $size_{i,t-1}$, and previous month return, $R_{i,t-1}$, in the regression to control for the size effect and bid-ask bounce (these are included as deviations from cross-sectional means to facilitate interpretation of the intercept).

We control for momentum by including the 52-week high momentum measures identified in George and Hwang (2004). These measures dominate others used in the literature in capturing momentum effects. Their definitions are as follows:

$52wkhW_{i,t-j}$ ($52wkhL_{i,t-j}$) equals one if $\dfrac{P_{i,t-j}}{high_{i,t-j}}$ is ranked among the top (bottom) $Q\%$ of all stocks in month t-j, and zero otherwise; where $P_{i,t-j}$ is the price of stock i at the end of month t-j and $high_{i,t-j}$ is highest month-end price of stock i during the 12-month period that ends on the last day of month t-j.

Estimates of the coefficient b_{0jt} can be interpreted as the return in month t to a "neutral" portfolio that was formed in month t-j that has hedged (zeroed out) the effects of deviations from average size and past return, momentum, and the effects of measure 1 and 2 dummies in predicting returns. The sum $b_{0jt} + b_{5jt}$ is the month-t return to a portfolio formed in month t-j that is long measure-1 winner stocks but that has hedged out all other effects. Consequently, b_{5jt} can be viewed as the return in excess of b_{0jt} earned by taking a long position j months ago in a "pure" measure-1 winner portfolio. The difference $b_{5jt} - b_{6jt}$ is the return to a "pure" zero investment portfolio that is formed by taking long positions in measure-1 winners and shorting an equal dollar amount of measure-1 losers j months ago. The remaining coefficients have similar interpretations (see Fama (1976)).

The *total* month-t returns involve portfolios formed over the prior 60 months. For a given measure, the total month-t return to "pure" measure-1 winner and loser portfolios

can be expressed as sums such as $S_{5t} = \frac{1}{60}\sum_{j=1}^{60} b_{5jt}$ and $S_{6t} = \frac{1}{60}\sum_{j=1}^{60} b_{6jt}$ where the individual

coefficients are computed from separate cross-sectional regressions for each $j = 1,...,60$.

Dividing by 60 rescales the sums to be in terms of monthly returns. The time-series

means of the month-by-month estimates of these sums, (e.g., $\overline{S}_5$ and $\overline{S}_6$) and associated

t-statistics, computed from the temporal distribution of sums, are reported in the tables.

Since we require five years of data to form portfolios and the shortest holding period we

analyze is one year, the first month for which the regression is estimated is January 1968.

In that month, 2223 stocks have valid data for return or market capitalization (size), but

only 1609 stocks have non-missing data for all of the variables (lagged return, size, five

year return, and five year low). The corresponding numbers increase to 7606 and 4856

by December 2001.

Table I is a correlation matrix for the indicator variables used in the regressions. These

variables identify whether or not a stock is a winner or loser for a particular month. Thus,

the correlations measure the average similarity of the cross sections of winners and losers

identified using the measures described above. A few observations are noteworthy. First,

there is a great deal of difference between the winners and losers formed using past five-

year return, *TR,* and nearness to historical highs and lows, *FYH* and *FYL*. The correlation

between the *TR* winner (loser) and *FYH* winner (*FYL* loser) variables is only 0.365

(0.469). This means that ranking stocks based on past return does not generally identify

stocks whose prices are near historical high or low points. Second, the *EWGO* winner

variable is highly correlated with the *EWGL* winner (0.797), *TR* winner (0.635) and *FYL*

15

winner (0.843) variables even though these variables differ substantially in their ability to explain reversals in the regressions below.

III. Results

A. Verification of Long-Term Reversals

We first verify the long-term reversal results in the existing literature using our regression approach. Winners and losers in DeBondt and Thaler (1985, 1987) are defined as the top and bottom 50 NYSE stocks on the CRSP tape ranked by past five-year return. Jegadeesh and Titman (2001) study returns at horizons up to five years for portfolios that include the top 10% and bottom 10% of stocks on CRSP ranked by past one-year return. Jegadeesh and Titman are interested in long-term reversals of momentum. They use last year's return to form portfolios because momentum is strongest at horizons of six months to one year. In this section, we use the traditional past five-year return, and 10% cutoffs as in Jegadeesh and Titman, to form winner and loser portfolios. The results are presented in Table II. Following Grinblatt and Moskowitz (2004), we report separate results with January included and January excluded. Most of our discussion focuses on returns earned over the entire five-year period, which appear in the last two columns in the table labeled (1,60). The other columns report returns over subintervals of the five-year horizon, with the exception of the fifth year to save space.

Our results are consistent with the earlier literature in documenting that there is a strong reversal. In fact, the magnitude of the reversal is greater in our results. In Jegadeesh and Titman's (2001) small-firm sample, the average reversal, as measured by the return to the

loser minus winner portfolio, is about 0.29% per month while ours is 0.56%

(0.30%+0.26%) per month over the five-year period labeled (1,60) in the table. This

figure is also larger than the 31.9% (over 60 months) reported by DeBondt and Thaler

(1987) with a more extreme cutoff—i.e., using only the top and bottom 50 NYSE stocks

on CRSP. Furthermore, the reversals we document begin in the first year after portfolio

formation (columns labeled (1,12)), while the reversals in Jegadeesh and Titman start in

the second year. Since we use a regression approach, these estimates capture the impact

of long-term reversals in isolation, controlling for the effects of short-term momentum,

size and bid-ask bounce.

Two other aspects of the results in Table II are worth noting. First, the strong loser

reversal comes exclusively from January. Outside January, there is no loser reversal at

any horizon. Over the five-year period after portfolio formation, the winner reversal

outside January is -0.32% per month, while the loser reversal is an insignificant 0.06%.

The fact that losers reverse only in January is similar to findings other attribute to tax loss

selling (see Roll (1983), Schultz (1985), D'Mello, Ferris and Hwang (2003), Grinblatt

and Keloharju (2004)). What is different here is that losers are defined with respect to

long-term returns, not just the return from the past year. Our finding that long-term

losers exhibit high January returns is consistent with that of Grinablatt and Moskowitz

(2004). They find that losers and consistent winners (identified based on the cumulative

return from months t-36 to month t-13) have January returns that are greater than returns

in other months. Their consistent long-term winner dummy has a significant positive

17

sign in January, which matches the smaller winner reversal in our results when the

January return is included. They attribute their findings to tax loss selling.[2]

Second, tax loss selling appears to affect the performance of winners as well, because

winner returns are mostly positive in January. For the entire five-year period and all

subintervals except the first, winner returns are more negative when January is excluded

than when January is included. This suggests that the redistribution of funds that results

from tax loss selling of losers puts upward price pressure on winners in January. Our

finding is consistent with Ritter (1988) who documents that, as a group, smalls stocks do

not display a noticeable December price decline though they register a large January

return. He advances a "parking the proceeds" hypothesis to supplement tax loss selling

in order to explain this. He cites discussions with brokers as indicating that investors

typically wait for several days or weeks before reinvesting the proceeds from their

December sales.[3]

Ritter (1988) also shows that when individuals reinvest, they disproportionately invest in

smaller stocks that are less liquid than larger stocks, which in turn amplifies the January

effect. Our finding that long-term winners and losers have higher January returns than

average stocks suggests that when investors finally do reinvest, they focus more on

winners and losers than on stocks in the middle 40% of long-term performance. This is

consistent with Ritter's findings. Smaller stocks are more volatile and are more likely to

become long-term winners or losers to begin with.

We also estimate regressions where the tax-related measures (numbered 4-6 above) are used in place of *TR* to classify stocks into winner and loser portfolios. The tables are omitted for brevity. All measures exhibit winner reversals at all horizons with nearly uniform significance. Similar to Table II, winner reversals are dampened in January relative to non-January months. Also apparent in those regressions is that winner reversals are largest for Five-Year Low and Equal-Weighted Gain Only. These are the measures the capital gains lock-in hypothesis predicts should exhibit reversals. Outside January, the winner reversals for *FYL* and *EWGO* are 0.43% and 0.41% per month respectively, which are both larger than the 0.35% for *EWGL* and the 0.32% for *TR* in Table II. Similar to *TR*, the loser reversals for *EWGL* that are significant are restricted to January. However, there are no loser reversals for *FYL* and *EWGO* even when January is included. These findings are consistent with tax loss selling as an explanation for *TR* and *EWGL* loser reversals, because portfolios formed using *FYL* and *EWGO* are based on embedded capital gains only and exclude losses.

Though Klein's (1999) model predicts lower returns to winner stocks relative to loser stocks, it does not follow that expected returns are negative for portfolios of winners. Portfolios of winners are exposed to systematic risk, and should therefore earn positive returns on average. This can be verified by adding the intercept to the coefficient on the individual winner dummies *TRW* (in Table II)*, FYLW, EWGLW*, and *EWGOW* (in results omitted)*. In all four cases, the sum of these coefficients is positive.

These results document that our regression approach captures clearly the long-term return

reversals documented in the literature that uses 10% or more extreme cutoffs in defining

winners and losers. To address the possibility that the results are unique to a narrow

definition of winners and losers, the remaining tests in the paper broaden the definition to

use 30% cutoffs. We replicate the tests described above using 30% cutoffs. Results for

TR are presented in Table III. The patterns and conclusions are the same except, not

surprisingly, the reversals are smaller in magnitude though still quite significant. In

Table III, the winner reversal for *TR* is a significant 0.16% per month outside January

compared to 0.32% per month using the 10% cutoff in Table II. In omitted results, the

winner reversal for *FYL* is a significant 0.25% per month outside January using the 30%

cutoff, compared to 0.43% per month using the 10% cutoff. The winner reversal for

EWGO is a significant 0.23% per month compared to 0.41% using the 10% cutoff.

The estimates above confirm the existence of reversals when winner and loser portfolios

are identified using the measures defined in the earlier section, and also support the

inferences of others that loser reversals are attributable to tax loss selling. To test

whether the overreaction or capital gains lock-in hypothesis is more descriptive of the

data, we estimate regressions that include multiple measures from the set of 1-6 above.

The coefficients in these regressions are average returns contributed by the long or short

side of a given strategy that hedges (zeros out) the contribution of other strategies whose

measures are also included in the regression. This enables us to test whether measures

based on the capital gains lock-in hypothesis dominate the other measures in explaining

reversals.

B. Tests of the Overreaction Hypothesis

As explained earlier, if reversals associated with past return winners are due to overreaction, then reversals should be stronger for winner stocks whose prices are near long-term highs than for winner stocks whose prices are not near long-term highs. We would expect that stocks classified as winners by both the traditional five-year return (*TR*) measure and the five year high (*FYH*) measure should have stronger reversals than stocks classified as winners by *TR* alone. In this case, the regression coefficient on an interaction between the *FYH* winner dummy and the *TR* winner dummy should be negative and significant. The estimates in Table IV indicate that this is not the case. In fact, the estimates are either insignificant or positive and significant, while the coefficients on the *TR* dummy remain significantly negative as in Table III where no interaction terms are included.

Consider the results in the last column. The coefficient of -0.21 on the *TR* winner dummy indicates that this group experiences a reversal of 0.21% per month relative to a portfolio that is neither a *TR* winner nor a *TR* loser. The coefficient on the interactive term indicates that *TR* winners whose prices are near a five-year high earn a significant 0.10% *more* per month than the return to all *TR* winners. These stocks reverse *less* at -0.11% (i.e., -0.21 + 0.10) per month relative to a portfolio that is neither a *TR* winner nor a *TR* loser. The same interpretation applies to estimates in other periods. None are consistent with the prediction of the overreaction hypothesis that *TR* winners whose prices are near a five-year high should have stronger reversals than *TR* winners whose

prices are not near a five-year high. In fact, the results suggest that at the one-year horizon, and overall, *TR* winners near *FYH* have significantly *weaker* reversals.

There are important differences across horizons. At the one-year horizon, the coefficient estimate on the winner interaction term is significant and its magnitude is the same or larger than that of the Five-Year Return Winner variable. Being near a long-term high *eliminates* the reversal that would otherwise occur at the one-year horizon. Things are different for years two through five. At those horizons, the interaction terms are small relative to the coefficient on the Five-Year Return variables and generally insignificant. In other words, being near a long-term high or low has *no impact* on winner reversals between years two and five. Consequently, for the five-year period as a whole, the reversal is mitigated if a stock is near a long-term high.

The key to understanding these differences probably relates to the findings in George and Hwang (2004). In particular, we believe the reversal is present even at the one-year horizon, but the momentum effect documented by George-Hwang offsets it at the one-year horizon for stocks near long-term highs. The winner interaction term identifies the best performers over the past five years whose prices are still near *five*-year highs. Even though we include the George-Hwang 52-week high indicator variables, they are unlikely to capture fully momentum effects associated with historical highs. The coefficient on our winner interaction partly reflects the momentum effect that is not captured fully by their variables. George-Hwang also document that their momentum effect does not reverse. The winner interaction effect at the one-year horizon does not reverse either,

which is why those stocks experience a mitigated reversal over the entire five-year horizon.

We now turn attention to *TR* losers. The results in Tables II and III indicate that *TR* losers do not reverse at any horizon when January returns are excluded. Those results suggest that loser reversals documented in the literature are likely due to tax loss selling rather than overreaction.   The results in Table IV support this. Table IV also examines whether *TR* losers reverse more strongly when their recent prices are near five-year lows—i.e., we test whether the regression coefficients on an interaction between the *TR* loser and *FYL* loser dummies are significantly positive. Although the estimates are positive for all horizons, they are small and none of them is significantly different from zero.

The evidence for both winners and losers is not consistent with the overreaction hypothesis as an explanation of long-term return reversals documented in the literature. The results are consistent with the view that loser reversals are due to tax loss selling. In the following section we will demonstrate that winner reversals can be explained by the capital gains lock-in hypothesis.

C. Tests of the Capital Gains Lock-In Hypothesis

The two predictions of the capital gains lock-in hypothesis are that (a) only winners reverse and (b) winners as identified by measures of embedded capital gains have stronger reversals than winners identified using other measures.

We construct four measures for tests of the capital gains lock-in hypothesis as explained earlier: Five-year low (*FYL*) and equal weighted gain only (*EWGO*) are the measures that predict winner reversals under the capital gains lock-in hypothesis. The traditional five-year return (*TR*) and equal-weighted gain and loss (*EWGL*) should not predict winner reversals in the presence of *FYL* or *EWGO* if the capital gains lock-in hypothesis is correct. We perform four pair-wise comparisons: (i) *FYL* vs. *TR*, (ii) *EWGO* vs. *TR*, (iii) *FYL* vs. *EWGL*, and (iv) *EWGO* vs. *EWGL*. If the capital gains lock-in hypothesis is correct, we expect *FYL* and *EWGO* to outperform *TR* and *EWGL* in predicting winner reversals, but not in predicting loser reversals. This is because *TR* and *EWGL* are computed in a manner that incorporates losses and should therefore exhibit reversals, but only in January relating to tax loss selling.

Table V Panel A compares *FYL* and *TR*.[4] Consistent with the capital gains lock-in hypothesis, *FYL* completely dominates *TR* at every horizon in predicting winner reversals. The magnitudes of *FYL* winner reversals are many times those corresponding to *TR*. The latter are insignificant at every horizon, which is in sharp contrast to the winner reversals in Tables II and III where *FYL* is not included in the regression. In the five year period after portfolio formation, *FYL* winner stocks earn -0.21% per month outside January. *TR* winners earn an insignificant -0.02%. The evidence for losers is also consistent with the capital gains lock-in hypothesis. There are no return reversals at any horizon for *FYL* loser portfolios, and significant reversals for *TR* losers are isolated in January.

It is interesting to note that the winner reversals measured by the *FYL* portfolio are greatest in the second year, with an excess return of -0.35% per month outside January, then decrease slowly. This is consistent with Klein's (1999) prediction that the premium caused by capital gains lock-in dissipates slowly. The reason reversals are not greatest in the first year may reflect the possibility that momentum is captured by the 52-week high dummies imperfectly. In this case, some momentum remains unexplained by the 52-week high dummies and our estimates understate the magnitude of the reversal.

It is also noteworthy that the positive January returns to winners documented in Tables II and III are present here also; *FYL* winners earn -0.14% when January is included versus -0.21% outside January. This is consistent with upward price pressure on winners to which December sale proceeds from losers are redistributed in January. However, this is not a specific prediction that emerges from Klein's (1999) model. He assumes that taxes are paid (or tax credits received) when the gain or loss is realized. Consequently, there is no seasonal concentration of gain/loss realization associated with the end of a tax year in his model.

Panel B of Table V compares *EWGO* and *TR*. The capital gains lock-in measure *EWGO* uniformly dominates *TR* in capturing winner reversals. Winner reversals of *EWGO* portfolios are slowly dissipating. *EWGO* losers do not exhibit reversals. The returns to *TR* loser portfolios exhibit significant reversals only if January is included. These results support capital gains lock-in as the explanation for long-term reversals.

We compare *FYL* and *EWGL* in Table VI Panel A, and *EWGO* and *EWGL* in Panel B. Recall that *EWGL* treats gains and losses symmetrically, while the capital gains lock-in hypothesis predicts that reversals should be associated with gains only. The results are similar to those of the regressions involving *TR* in Table V, providing additional evidence in favor of the capital gains lock-in explanation for long-term return reversals.

The capital gain lock-in motivated winner dummies in the (1,60) window are not significantly negative at the 5% level in some columns of Tables V and VI when January returns are included (though most are significant at the 10% level) . As discussed in connection with the results of Table II, these coefficients are probably contaminated by tax loss selling and possible "parking the proceeds" effects.[5] One could argue that these estimates should be discounted in drawing conclusions regarding both the overreaction hypothesis and the capital gain lock-in hypothesis. However, it turns out that the contamination is not so large that it affects the results if winners and losers are defined in accordance with the stricter criteria used in the earlier literature. If we tighten our criteria, *t*-statistics are above two whether January is excluded or not, and the non-tax-motivated measures remain insignificant.[6]

To examine whether our conclusions are attributable to differences in the risk of the strategies implied by the measures we use, we re-examine Table V using risk-adjusted returns. Each coefficient reported in Table V is a time-series average of sums (e.g., S_{5t} and S_{6t} defined above) of monthly raw returns to a particular portfolio strategy, where each monthly raw return is obtained from a cross-sectional regression. To compute the risk-adjusted return to a particular strategy, we estimate the intercept of a time-series

26

regression of that strategy's sums on the Fama-French (1993) factor realizations.[7] These

intercepts are risk-adjusted returns to the "pure" portfolios described in Section II. The

risk-adjusted return estimates are reported in Table VII along with their regression *t*-

statistics.

In both panels of Table VII, the tax-based measures dominate *TR* in predicting winner

reversals—the tax based measures are significant at almost all horizons whether January

is included or not, and the *TR* winner dummy is uniformly insignificant. The significance

of *FYL* and *EWGO* is stronger in risk-adjusted returns than in raw returns, with t-statistics

greater than three for the 60-month results both with and without January. The *TR* loser

dummy is mostly insignificant and, as in Table V, it is significant only if January is

included. However, its significance is less than with raw returns because the Fama-

French size factor captures the component of returns associated with tax loss selling.

We now briefly shift perspective from testing hypotheses about why returns reverse to

that of how investors with neutral tax positions might benefit from an awareness of these

reversals. Tables V – VII indicate that *FYL* has a slight numerical advantage over *EWGO*

of about two or three basis points per month.   It is also computationally easier to assess

as investors need only rank stocks by *FYL*.   Though not reported in tables, we estimated

regressions using 10% cutoffs of *FYL* (i.e., similar to those in Table II, but using *FYL*

instead of *TR*). Investors could have avoided an average raw return loss of 0.33%

(0.43%) per month (outside January) by avoiding stocks ranked in the top *FYL* decile

when other close substitute stocks are available. We also reproduced the regressions

using risk-adjusted returns. Investors could have avoided a risk-adjusted loss of 0.30% (0.36%) per month (outside January) by avoiding top *FYL* decile stocks. Both raw and risk-adjusted estimates are highly statistically significant.

D. Comparison of our Risk-Adjusted Results to the Existing Literature

Various authors such as Chan (1988) and Ball and Kothari (1989) argue that the profit to strategies that exploit long term reversals is primarily compensation for systematic risk. Ball and Kothari use a technique developed by Ibbotson (1975) to estimate the betas of winner and loser portfolios that allows the portfolio betas to vary across years relative to the portfolio formation year. This enables them to show that loser (winner) portfolio betas increase (decrease) between the ranking and performance periods. Using the Sharpe-Lintner CAPM as the benchmark, they then show that this shift in betas is large enough that abnormal returns to DeBondt and Thaler's contrarian strategy disappear. Chopra, Lakonishok and Ritter (1992) draw the opposite conclusion from a similar analysis. The difference is that they *estimate* the price of beta risk rather than assuming it is the difference between the average market return and the risk free rate. They find that the Sharpe-Lintner price of beta risk is double their empirical estimate, and that the profit to DeBondt and Thaler's contrarian strategy remains significant when returns are adjusted for risk using their estimate.

DeBondt and Thaler and the empirical work cited in the previous paragraph all form portfolios in December. Ball, Kothari and Shanken (1995) observe that if profit to contrarian strategies is attributable to overreaction, then the profitability of these

strategies should not depend on when portfolios are formed. Thus, Ball, Kothari and

Shanken compare the performance of portfolios formed in June to portfolios formed in

December. They find that returns to loser portfolios formed in June are much less than

those formed in December, and that risk adjusted returns to a contrarian strategy based on

June portfolios are actually negative on average. Fama and French (1996) examine

monthly returns to portfolios consisting of past winner and loser stocks—in effect,

forming contrarian portfolios monthly rather than yearly or twice yearly, but holding the

portfolios for a single month. They then examine the intercepts of time-series regressions

of the returns to these portfolios on the Fama-French (1993) factors. They find that the

Fama-French (1993) factors explain the one-month-ahead returns to long-term (five-year)

winners and losers, but not the one-month-ahead returns to short-term (one-year) winners

and losers (i.e., the model does not explain short-term momentum).

These studies analyze NYSE stocks only, with the exception of Ball, Kothari and

Shanken, which includes AMEX stocks. These studies also define winners/losers as the

top/bottom 50 performers during the ranking period, or a close approximation such as

top/bottom 5% of NYSE stocks. The exception is Fama and French (1996) who use the

top and bottom 10% of NYSE stocks.

The sample period examined in this paper is longer and the sample itself is broader.

Fama and French (1996), for example, use only NYSE stocks from 1963-1993, while we

use all CRSP stocks including NASDAQ from 1963-2001. In addition, we form

portfolios monthly and our definition of winner and loser stocks is more inclusive at 30%

than even the 10% cutoffs used by Fama and French (1996). Our risk adjustment is based on the Fama-French (1993) factors, and is therefore similar to Fama and French (1996). However, the most notable difference is that we estimate Fama and MacBeth (1973) style regressions, which enable us to draw sharper inferences. We are able to hedge out the effects of size, momentum and bid-ask bounce to isolate the monthly return attributable only to whether a stock is a winner or loser, has a large or small embedded capital gain, etc. This, and the fact that the entire sample of stocks (not just winners and losers) is used to assess the significance of returns to the portfolios of interest, yield more powerful tests than those based on subsamples of equal-weighted winner and loser portfolios. This is apparent even in comparing Tables II and III where we confirm DeBondt and Thaler's findings using a more inclusive definition of winners and losers.

The evidence in this paper leads to different conclusions about why reversals exist. The other studies find their results to be driven by reversals of loser stocks. We show that loser reversals are due to the effect of tax loss selling on January returns, and that winners experience significant reversals in raw and risk-adjusted returns. In fact, winner reversals are stronger in risk-adjusted returns than in raw returns. We show that winner reversals in both raw and risk-adjusted returns are attributable to embedded capital gains rather than past performance.[8]

In diagnostic tests (tables omitted for brevity), we investigate why the risk-adjusted results for winners are actually more significant than those based on raw returns. This is because the loadings of winner portfolios are negative on the HML (book-to-market) factor, and positive on both SMB (size) and MKT (market) factors. This means that,

compared with a neutral portfolio, the winner portfolios tend to be more heavily weighted toward smaller firms (this makes sense because smaller firms are more volatile and more likely to be winners or losers), have lower book to market ratios (because as winners, market value has increased and book to market is smaller), and have higher market betas. The effects of SMB and MKT dominate that of HML such that after adjusting for risk, risk adjusted returns are more negative than raw returns.

Finally, it is worth noting that the intercepts for the regressions in Table VII are risk-adjusted returns to "neutral" portfolios (i.e., portfolios that hedge out the effects of bid-ask bounce, size, momentum, long-term return and the tax variables). Though not reported, these estimates are insignificant across all horizons, indicating that the Fama-French factors do a good job explaining the returns to neutral portfolios. This is reassuring because it suggests that our risk adjustment procedure is well specified, and that the larger negative risk-adjusted versus raw return for winners is not an artifact of using a misspecified model of expected returns.

E. Evidence from a Regime without Taxes

Our evidence so far favors the capital gains lock-in hypothesis over the overreaction hypothesis as an explanation for long-term reversals in US stock market data. The approach that led to this conclusion was one of comparing the predictive power of measures of past performance that relate to one or the other of these hypotheses using data in which either explanation could be true. In this section, we use a different approach. Here we examine data for which the capital gains lock-in hypothesis *cannot* be true. Hong Kong does not tax investment income—neither capital gains nor dividends.

If the capital gains lock-in hypothesis is the correct explanation for reversals in the US data, then reversals should not exist in Hong Kong data. Alternatively, if overreaction is the correct explanation, reversals should exist in Hong Kong data, and be explained by measures of extreme past price changes, especially when the prices of winners (losers) are near long-term highs (lows).

Data for Hong Kong stocks are obtained from the Pacific Basin Capital Market (PACAP) database compiled at the University of Rhode Island. The PACAP database available to us covers the period from January 1980 through December 2000. Since we require five-years of data to form portfolios, the first month in which we run the Fama-MacBeth regressions is January 1985. The numbers of stocks that have return or size data in January 1985 is 253, but only 61 have non-missing data on all the variables required for the regression. The corresponding numbers increase to 744 and 334 by December 2000. The potential impact of the difference in sample sizes between US and Hong Kong data is addressed below.

E.1 Tests of the Overreaction Hypothesis

Table VIII reports tests of the overreaction hypothesis using returns data from Hong Kong. These tests correspond to those reported for US stocks in Table IV. The analysis examines whether reversals occur for five-year winner and loser stocks, and whether winners (losers) whose prices are near five-year highs (lows) exhibit stronger reversals than those whose prices are not near five-year highs (lows).

The results do not support the overreaction hypothesis. The uniform insignificance of the five-year return (*TR*) winner and loser dummies is consistent with the hypothesis that there are no long-term reversals in Hong Kong data. This is very different from the results in Table IV for US data. The coefficients on the interactive terms are either insignificant or positive, which is also inconsistent with the overreaction hypothesis. This reinforces our earlier conclusion that the winner reversal predicted by *TR* in the US market is not explained by overreaction.

Analyzing US data in Table IV shows that *TR* losers reverse only when January is included. We concluded there that loser reversals are due to the tax loss selling induced January effect, and not to overreaction. The results in Table VIII lend further weight to this conclusion because there are no loser reversals in Hong Kong even when January is included. The absence of taxes on investment income in Hong Kong implies that there should be no evidence of a tax loss selling induced January effect.

E.2 Tests of the Capital Gains Lock-In Hypothesis

FYL winners are stocks with large embedded capital gains. If the capital gains lock-in hypothesis is the explanation for the reversals we found in raw and risk-adjusted US stock returns in Tables V and VII, then such reversals should be absent in a market without taxes. Indeed, Table IX Panel A documents that *FYL* winners do not exhibit reversals in raw Hong Kong returns. Panel B reports results where *EWGO* is used instead of *FYL* as the measure of embedded capital gains. The inferences are almost identical to those for *FYL*. The only difference is that at horizon (37,48), *TR* winners exhibit

significant *continuations* when January is included. Despite this, there is no horizon at which *EWGO* winners in Hong Kong exhibit return reversals as documented for US stocks. The ability of the *FYL* and *EWGO* winner dummies to predict reversals in US returns, and the inability of these dummies to predict reversals in Hong Kong returns, are consistent with the hypothesis that capital gains lock-in is the source of the winner reversals we observe in US stock returns.

In unreported results, we also examined risk-adjusted returns to Hong Kong stocks using *FYL* and *EWGO* dummies.   The inferences mirror those of Table IX in that there is no evidence of significant reversals of *FYL* or *EWGO* winners. The only difference is that the continuation documented in raw returns for *TR* winners at the (37,48) horizon in the *EWGO* regression is not present in risk-adjusted returns.[9]

E.3 Robustness

The inferences in all these tables are drawn from distributions of monthly coefficient estimates. Since the Hong Kong sample is much shorter than the US sample, the number of monthly observations is much smaller for Hong Kong, and it is possible that this is driving the difference in results between the two samples. Moreover, if smaller cross sections yield noisier regression coefficient estimates that form the time series from which we draw inferences, then the tests that use the smaller Hong Kong cross sections will have low power relative to the tests that use US data.

To address this possibility, we conduct Hotelling's T^2 tests of the null that differences

between monthly regression coefficient vectors for the Hong Kong and US samples are

drawn from a distribution with a zero mean vector. We do this for the (1,60) regressions

involving *FYL* and *EWGO* with January included and January excluded. These tests limit

the US sample to match the time dimension of the Hong Kong sample, and account for

the difference in the precisions of US versus HK estimates that might result from

differences in the sizes of the cross sections.[10] For each of the four regressions, we

conduct the Hotelling test for the entire vector of coefficients, and separately for the

subvector of coefficients relating to only to the capital gains lock-in hypothesis

(specifically, the FYL or EWGO dummy, and TR winner and loser dummies). All eight

tests reject the null at the 5% level of significance. In other words, the differences in

coefficient estimates between the two countries' documented in the tables are not likely

to be attributable to sampling error.

IV. Conclusion

In this paper, we test the leading behavioral hypothesis for why stock returns exhibit

long-term reversals against an explanation based on a rational model of tax avoidance.

The overreaction hypothesis postulates that long-term reversals occur because investors

overreact to news, and returns reverse as investors come to realize they overreacted.

The peak of overreaction to good news is likely to occur when stocks have reached a

long-term high, indicating that good news has arrived, investors have overreacted, and

the reversal has not yet occurred. Similarly, the peak of overreaction to bad news is

likely to occur when prices are near long-term lows. Thus, the overreaction hypothesis

predicts that reversals should be stronger for winner (loser) stocks whose prices are near long-term highs (lows) than winners (losers) whose prices are not near long-term highs (lows). We find no evidence that winner and loser stocks that are near long-term highs and lows exhibit stronger reversals. In fact, when nearness to long-term highs or lows is statistically significant, the reversals are actually weaker.

We also examine a tax-based explanation of reversals. Since capital gains are taxed upon realization, the reservation prices of rational investors for selling stocks with embedded capital gains are elevated. As shares eventually turn over to buyers without large embedded gains, the marginal investor's reservation price falls resulting in a slowly dissipating reversal. This intuition is formalized in a general equilibrium framework by Klein (1999). This does not apply to stocks with embedded losses, however, so loser stocks should not experience reversals.

We find that measures defined in accordance with the capital gains lock-in hypothesis subsume the usual past return measure in predicting winner reversals, and that winner reversals are indeed slowly dissipating. Moreover, we find no loser reversals outside of January. This is inconsistent with the overreaction hypothesis, which predicts reversals for losers both in and outside January. Instead, the evidence suggests that loser reversals are attributable to tax loss selling in December.

Finally, we examine stock returns data from Hong Kong where investment income is not taxed (neither dividends nor capital gains). Long-term reversals do not exist for winners

or losers, in or outside January. Like our other findings, this is inconsistent with the overreaction hypothesis and is jointly consistent with the hypotheses that the reversals we document for winners in the US data are attributable to capital gains lock-in, and that the reversals for US losers in January are attributable to tax loss selling.

Rational explanations for long-term reversals that are not based on taxes are presented in Berk, Green and Naik (1999), Lewellen and Shanken (2002) and Brav and Heaton (2002). Berk, Green and Naik model the dynamics of project adoption and disposition by firms. When new low-risk projects are unexpectedly adopted, the current stock price increases and firm risk decreases. The lower average future returns that follow create the appearance of reversals. The opposite happens when low-risk projects die off. Thus, reversals can appear because, as assets turn over, good (bad) news is associated with reductions (increases) in the future riskiness of the firm's projects. Brav and Heaton and Lewellen and Shanken are models of rational learning. They point out that investors acting in real time do not know the parameters of the data generating process for security returns. Even if investors learn as perfect Bayesians, they place heavy weight on extreme observations that are simply outliers because they have no way to know in real time that such observations are outliers rather than indications that a shift in the data generating process has occurred. This generates overreaction in beliefs to both good and bad news that, in turn, affects market prices. Reversals occur as investors learn that they overreacted in the past. While all three models predict the existence of long-term reversals, none predicts why tax-based variables would capture reversals better than past return measures or the asymmetric seasonal pattern in reversals that we and others have

documented—that loser reversals occur only in January. These models also do not predict differences in the incidence of reversals across countries that we document here.

Conrad, Cooper and Kaul (2003) explore whether data snooping can explain the strength with which past returns (and other variables) predict future returns. They examine profits to strategies that take long (short) positions in winners (losers) determined by "predictors" that are simulated random series, but chosen from among a larger set of random series because their realizations just happen to predict returns. Conrad et.al. show that in some cases, the profits to these strategies are similar to those documented in the literature using economic variables as predictors. They show that the problem is more severe the tighter the cutoffs (i.e., top and bottom 10% versus 30%), and the greater the number of predictors used jointly to sort stocks into cells that determine winners and losers.

Even studies that test specific theories cannot avoid data snooping entirely because theory is written partly to explain stylized facts of which the theorist is aware. However, in designing our tests, we have taken several steps to minimize the impact of data snooping on our inferences. First, in light of Conrad et.al.'s findings, we use generous cutoffs. Second, the variables we use are new to this literature and designed to capture, in as focused a manner as possible, the predictions of the specific hypotheses of overreaction and tax avoidance, both of which are predicted by formal models. One of Klein's (1999) modeling objectives may have been to explain long-term reversals, so it is possible that the tax-motivated variables are significant by chance. However, (i) they are consistent

with a model that makes *asymmetric* predictions about their significance, (ii) they

dominate variables that have been used for years in the literature, and (iii) they fail to

have explanatory power in a regime without taxes. It seems unlikely to us that this

combination of findings is a result of data snooping.

Table I – Correlation Matrix, 30% Cutoff

Using monthly data from January 1968 and December 2001, we construct indicator variables for the 52-week high momentum variable in George and Hwang (2004) and for each of the measures described in the text. $52wkhW_{i,t}$ ($52wkhL_{i,t}$) is the 52-week high winner (loser) dummy that takes the value of 1 if the 52-week high measure for stock i is ranked in the top (bottom) 30% in month t, and zero otherwise. The 52-week high measure in month t is the ratio of price level in month t to the maximum price achieved in months t-12 to t. Similar to the construction of the 52-week high winner and loser dummies, the other variables (e.g., *FYH* Winner and *FYL* Loser) are winner and loser dummies defined with respect to their measures (e.g., *FYH* and *FYL*) using a 30% cutoff. $FYL_{i,t}$ is the price of stock i at time t relative to its five year low over the period from t-60 to t. $FYH_{i,t}$ is the price of stock i at time t relative to its five year high over the period from t-60 to t. $TR_{i,t}$ is the five-year return for stock i over the period from t-60 to t. $EWGO_{i,t}$ is the gain only, excluding losses, embedded in stock i on date t under the assumption that the shares are acquired uniformly over the period from t-60 to t. $EWGL_{i,t}$ is the gain net of losses embedded in stock i on date t under the assumption that the shares are acquired uniformly over the period from t-60 to t. Numbers reported in the table are time-series averages of cross-sectional correlations.

	FYH Loser	*FYH* Winner	*FYL* Loser	*FYL* Winner	*52-Wk High* Loser	*52-Wk High* Winner	*TR* Loser	*TR* Winner	*EWGO* Loser	*EWGO* Winner	*EWGL* Loser	*EWGL* Winner
FYH Loser	1.000											
FYH Winner	-0.429	1.000										
FYL Loser	0.308	-0.284	1.000									
FYL Winner	-0.216	0.261	-0.430	1.000								
52-Wk High Loser	0.534	-0.395	0.242	-0.132	1.000							
52-Wk High Winner	-0.314	0.543	-0.224	0.171	-0.434	1.000						
TR Loser	0.564	-0.391	0.469	-0.312	0.289	-0.212	1.000					
TR Winner	-0.288	0.365	-0.400	0.633	-0.150	0.167	-0.429	1.000				
EWGO Loser	0.285	-0.282	0.868	-0.432	0.215	-0.214	0.458	-0.402	1.000			
EWGO Winner	-0.218	0.284	-0.429	0.843	-0.127	0.173	-0.320	0.635	-0.432	1.000		
EWGL Loser	0.629	-0.417	0.617	-0.399	0.424	-0.324	0.654	-0.401	0.597	-0.391	1.000	
EWGL Winner	-0.341	0.452	-0.426	0.741	-0.219	0.270	-0.398	0.677	-0.426	0.797	-0.429	1.000

Table II - Traditional Measure, 10% Cutoff

Each month between January 1968 and December 2001, 60 ($j=1,…,60$) cross-sectional regressions of the following form are estimated:

$$R_{it} = b_{0jt} + b_{1jt} R_{i,t-1} + b_{2jt}\, size_{i,t-1} + b_{3jt}\, 52wkhW_{i,t-j} + b_{4jt}\, 52wkhL_{i,t-j} + b_{5jt}\, TRW_{i,t-j} + b_{6jt}\, TRL_{i,t-j} + e_{ijt}$$

where R_{it} is the return to stock i in month t, $R_{i,t-1}$ and $size_{i,t-1}$ are the return and the natural logarithm of market capitalization of stock i in month t-1 net of the month $t-1$ cross-sectional mean; $52wkhW_{i,t-j}$ ($52wkhL_{i,t-j}$) is the 52-week high winner (loser) dummy that takes the value of 1 if the 52-week high measure for stock i is ranked in the top (bottom) 10% in month $t-j$, and zero otherwise. The 52-week high measure in month $t-j$ is the ratio of price level in month $t-j$ to the maximum price achieved in months $t-j-12$ to $t-j$. Similar to the construction of the 52-week high winner and loser dummies, the measures TRW and TRL are winner and loser dummies defined with respect to TR using a 10% cutoff. $TR_{i,t-j}$ is the five-year return for stock i over the period from $t-j-60$ to $t-j$. The coefficient estimates of a given independent variable are averaged over $j=1,…,12$ for columns labeled (1,12), $j=13,…,24$, for columns labeled (13,24),…, $j=1,…,60$ for columns labeled (1,60). The numbers reported in the table are the time-series averages of these averages. They are in percent per month. The accompanying t-statistics are calculated from the time series.

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Intercept	1.25	0.85	1.27	0.85	1.39	0.97	1.41	0.97	1.40	0.98
	(4.85)	(3.43)	(4.88)	(3.37)	(5.32)	(3.85)	(5.44)	(3.93)	(5.32)	(3.89)
$R_{i,t-1}$	-6.82	-5.83	-6.81	-5.65	-6.82	-5.62	-6.73	-5.51	-6.62	-5.43
	(-15.04)	(-14.47)	(-13.83)	(-13.04)	(-13.50)	(-12.68)	(-13.20)	(-12.28)	(-12.94)	(-12.21)
Size	-0.10	0.02	-0.06	0.07	-0.09	0.05	-0.10	0.04	-0.10	0.02
	(-2.48)	(0.42)	(-1.41)	(1.86)	(-2.05)	(1.24)	(-2.26)	(0.97)	(-2.39)	(0.60)
52-Wk High Winner	0.32	0.43	-0.14	-0.05	-0.08	-0.02	-0.09	-0.06	-0.01	0.05
	(4.76)	(6.52)	(-2.87)	(-1.10)	(-1.75)	(-0.37)	(-2.11)	(-1.48)	(-0.28)	(2.01)
52-Wk High Loser	-0.40	-0.96	0.19	-0.21	-0.01	-0.28	0.09	-0.16	0.02	-0.34
	(-2.02)	(-6.09)	(1.21)	(-1.64)	(-0.04)	(-2.03)	(0.60)	(-1.13)	(0.12)	(-2.78)
Five-Year Return Winner	-0.33	-0.27	-0.39	-0.42	-0.29	-0.37	-0.17	-0.28	-0.26	-0.32
	(-2.61)	(-2.09)	(-3.62)	(-3.77)	(-2.84)	(-3.61)	(-1.57)	(-2.66)	(-2.58)	(-3.11)
Five-Year Return Loser	0.66	0.19	0.43	0.15	0.13	-0.06	0.13	0.02	0.30	0.06
	(3.99)	(1.37)	(2.97)	(1.13)	(1.05)	(-0.49)	(1.20)	(0.19)	(2.61)	(0.60)

41

Table III - Traditional Measure, 30% Cutoff

Each month between January 1968 and December 2001, 60 ($j=1,\ldots,60$) cross-sectional regressions of the following form are estimated:

$$R_{it} = b_{0jt} + b_{1jt}R_{i,t-1} + b_{2jt}\,size_{i,t-1} + b_{3jt}\,52wkhW_{i,t-j} + b_{4jt}\,52wkhL_{i,t-j} + b_{5jt}\,TRW_{i,t-j} + b_{6jt}\,TRL_{i,t-j} + e_{ijt}$$

where R_{it} is the return to stock i in month t, $R_{i,t-1}$ and $size_{i,t-1}$ are the return and the natural logarithm of market capitalization of stock i in month t-1 net of the month t-1 cross-sectional mean; $52wkhW_{i,t-j}$ ($52wkhL_{i,t-j}$) is the 52-week high winner (loser) dummy that takes the value of 1 if the 52-week high measure for stock i is ranked in the top (bottom) 30% in month t-j, and zero otherwise. The 52-week high measure in month t-j is the ratio of price level in month t-j to the maximum price achieved in months t-j-12 to t-j. Similar to the construction of the 52-week high winner and loser dummies, the measures TRW and TRL are winner and loser dummies defined with respect to TR using a 30% cutoff. $TR_{i,t-j}$ is the five-year return for stock i over the period from t-j-60 to t-j. The coefficient estimates of a given independent variable are averaged over $j=1,\ldots,12$ for columns labeled (1,12), $j=13,\ldots,24$, for columns labeled (13,24),…, $j=1,\ldots,60$ for columns labeled (1,60). The numbers reported in the table are the time-series averages of these averages. They are in percent per month. The accompanying t-statistics are calculated from the time series.

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Intercept	1.28	0.94	1.26	0.89	1.37	0.98	1.40	1.00	1.39	1.02
	(5.25)	(3.99)	(5.10)	(3.73)	(5.51)	(4.12)	(5.61)	(4.20)	(5.56)	(4.24)
$R_{i,t-1}$	-6.93	-5.93	-6.89	-5.71	-6.91	-5.70	-6.80	-5.57	-6.70	-5.49
	(-15.42)	(-14.94)	(-13.97)	(-13.25)	(-13.69)	(-12.94)	(-13.44)	(-12.58)	(-13.12)	(-12.45)
Size	-0.13	-0.02	-0.06	0.06	-0.09	0.04	-0.10	0.03	-0.11	0.01
	(-3.25)	(-0.52)	(-1.47)	(1.67)	(-2.07)	(1.15)	(-2.36)	(0.86)	(-2.59)	(0.32)
52-Wk High Winner	0.25	0.31	-0.07	-0.01	-0.02	0.02	-0.03	-0.02	0.02	0.06
	(4.58)	(5.75)	(-1.86)	(-0.37)	(-0.62)	(0.53)	(-0.72)	(-0.47)	(0.90)	(2.29)
52-Wk High Loser	-0.41	-0.74	0.07	-0.20	0.00	-0.20	0.02	-0.15	-0.05	-0.28
	(-3.43)	(-7.31)	(0.70)	(-2.21)	(0.01)	(-2.06)	(0.18)	(-1.51)	(-0.51)	(-3.12)
Five-Year Return Winner	-0.17	-0.13	-0.19	-0.21	-0.13	-0.18	-0.10	-0.17	-0.12	-0.16
	(-2.19)	(-1.66)	(-2.74)	(-2.99)	(-1.92)	(-2.71)	(-1.49)	(-2.55)	(-1.97)	(-2.54)
Five-Year Return Loser	0.31	0.07	0.21	0.06	0.12	0.04	0.10	0.06	0.18	0.06
	(3.32)	(0.82)	(2.64)	(0.83)	(1.81)	(0.57)	(1.75)	(0.90)	(3.05)	(1.14)

Table IV – Traditional Measure and Its Interactions with Five-Year High and Five-Year Low

Each month between January 1968 and December 2001, 60 ($j=1,…,60$) cross-sectional regressions of the following form are estimated:

$$R_{it} = b_{0jt} + b_{1jt} R_{i,t-1} + b_{2jt}\, size_{i,t-1} + b_{3jt}\, 52wkhW_{i,t-j} + b_{4jt}\, 52wkhL_{i,t-j} + b_{5jt}\, TRW_{i,t-j} + b_{6jt}\, TRL_{i,t-j} + b_{7jt}\, TRW_{i,t-j}*FYHW_{i,t-j} + b_{8jt}\, TRL_{i,t-j}*FYLL_{i,t-j} + e_{ijt}$$

where R_{it} is the return to stock i in month t, $R_{i,t-1}$ and $size_{i,t-1}$ are the return and the natural logarithm of market capitalization of stock i in month $t-1$ net of the month $t-1$ cross-sectional mean; $52wkhW_{i,t-j}$ ($52wkhL_{i,t-j}$) is the 52-week high winner (loser) dummy that takes the value of 1 if the 52-week high measure for stock i is ranked in the top (bottom) 30% in month $t-j$, and zero otherwise. The 52-week high measure in month $t-j$ is the ratio of price level in month $t-j$ to the maximum price achieved in months $t-j-12$ to $t-j$. Similar to the construction of the 52-week high winner and loser dummies, *FYLL* is the loser dummy defined with respect to *FYL* using a 30% cutoff. $FYL_{i,t-j}$ is the price of stock i at time $t-j$ relative to its five year low over the period from $t-j-60$ to $t-j$. *FYHW* is the winner dummy defined with respect to *FYH* using a 30% cutoff. $FYH_{i,t-j}$ is the price of stock i at time $t-j$ relative to its five year high over the period from $t-j-60$ to $t-j$. The measures *TRW* and *TRL* are winner and loser dummies defined with respect to *TR* using a 30% cutoff. $TR_{i,t-j}$ is the five-year return for stock i over the period from $t-j-60$ to $t-j$. The coefficient estimates of a given independent variable are averaged over $j=1,…,12$ for columns labeled (1,12), $j=13,…,24$, for columns labeled (13,24),…, $j=1,…,60$ for columns labeled (1,60). The numbers reported in the table are the time-series averages of these averages. They are in percent per month. The accompanying t-statistics are calculated from the time series.

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Intercept	1.30	0.97	1.27	0.89	1.37	0.99	1.40	1.01	1.40	1.02
	(5.35)	(4.10)	(5.14)	(3.76)	(5.52)	(4.14)	(5.63)	(4.22)	(5.60)	(4.28)
$R_{i,t-1}$	-6.96	-5.96	-6.94	-5.75	-6.96	-5.75	-6.84	-5.61	-6.73	-5.52
	(-15.60)	(-15.15)	(-14.16)	(-13.47)	(-13.88)	(-13.18)	(-13.56)	(-12.72)	(-13.26)	(-12.62)
Size	-0.13	-0.02	-0.06	0.06	-0.09	0.04	-0.10	0.03	-0.11	0.01
	(-3.33)	(-0.60)	(-1.50)	(1.65)	(-2.07)	(1.19)	(-2.37)	(0.85)	(-2.62)	(0.30)
52-Wk High Winner	0.18	0.24	-0.09	-0.03	-0.02	0.02	-0.04	-0.03	0.00	0.04
	(3.69)	(4.82)	(-2.36)	(-0.90)	(-0.59)	(0.53)	(-1.15)	(-0.82)	(0.13)	(1.52)
52-Wk High Loser	-0.40	-0.72	0.06	-0.20	-0.01	-0.21	0.02	-0.15	-0.05	-0.28
	(-3.23)	(-6.92)	(0.58)	(-2.18)	(-0.11)	(-2.10)	(0.19)	(-1.51)	(-0.48)	(-2.99)
Five-Year Return Winner	-0.35	-0.33	-0.21	-0.25	-0.10	-0.15	-0.13	-0.19	-0.17	-0.21
	(-4.43)	(-4.09)	(-2.98)	(-3.45)	(-1.35)	(-2.04)	(-1.76)	(-2.60)	(-2.51)	(-3.10)
Five-Year Return Loser	0.27	0.04	0.11	-0.02	0.06	-0.04	0.05	-0.04	0.12	-0.02
	(2.39)	(0.35)	(1.02)	(-0.21)	(0.56)	(-0.36)	(0.59)	(-0.46)	(1.35)	(-0.20)
Five-Year Return Winner* Five-Year High Winner	0.35	0.39	0.08	0.11	-0.03	-0.02	0.06	0.05	0.09	0.10
	(6.20)	(6.93)	(1.42)	(1.86)	(-0.54)	(-0.33)	(1.03)	(0.85)	(2.55)	(2.66)
Five-Year Return Loser* Five-Year Low Loser	0.04	0.05	0.15	0.13	0.08	0.10	0.05	0.12	0.07	0.11
	(0.44)	(0.47)	(1.56)	(1.32)	(0.81)	(0.94)	(0.53)	(1.25)	(0.82)	(1.26)

Table V – Traditional Measure and Measures of Embedded Gains

Each month between January 1968 and December 2001, 60 ($j=1,...,60$) cross-sectional regressions are estimated where R_{it} is the return to stock i in month t, $R_{i,t-1}$ and $size_{i,t-1}$ are the return and the natural logarithm of market capitalization of stock i in month $t-1$ net of the month $t-1$ cross-sectional mean; $52wkhW_{i,t-j}$ ($52wkhL_{i,t-j}$) is the 52-week high winner (loser) dummy that takes the value of 1 if the 52-week high measure for stock i is ranked in the top (bottom) 30% in month $t-j$, and zero otherwise. The 52-week high measure in month $t-j$ is the ratio of price level in month $t-j$ to the maximum price achieved in months $t-j$-12 to $t-j$. Similar to the construction of the 52-week high winner and loser dummies, the measures TRW and TRL are winner and loser dummies defined with respect to TR using a 30% cutoff. $TR_{i,t-j}$ is the five-year return for stock i over the period from $t-j$-60 to $t-j$. The measures $FYLW$ and $FYLL$ in Panel A ($EWGOW$ and $EWGOL$ in Panel B) are winner and loser dummies defined with respect to FYL ($EWGO$) using a 30% cutoff. $FYL_{i,t-j}$ is the price of stock i at time $t-j$ relative to its five year low over the period from $t-j$-60 to $t-j$. $EWGO_{i,t-j}$ is the gain only, excluding losses, embedded in stock i on date $t-j$ under the assumption that the shares are acquired uniformly over the period from $t-j$-60 to $t-j$. The coefficient estimates of a given independent variable are averaged over $j=1,...,12$ for columns labeled (1,12), $j=13,...,25$, for columns labeled (13,24),..., $j=1,...,60$ for columns labeled (1,60). The numbers reported in the table are the time-series averages of these averages. They are in percent per month. The accompanying t-statistics are calculated from the time series. Coefficients on control variables are omitted for brevity.

Panel A

$$R_{it} = b_{0jt} + b_{1jt} R_{i,t-1} + b_{2jt} size_{i,t-1} + b_{3jt} 52wkhW_{i,t-j} + b_{4jt} 52wkhL_{i,t-j} + b_{5jt} TRW_{i,t-j} + b_{6jt} TRL_{i,t-j} + b_{7jt} FYLW_{i,t-j} + b_{8jt} FYLL_{i,t-j} + e_{ijt}$$

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Five-Year Return Winner	-0.08 (-1.71)	-0.02 (-0.49)	0.00 (-0.01)	0.02 (0.50)	-0.01 (-0.28)	-0.01 (-0.14)	-0.04 (-0.86)	-0.03 (-0.64)	-0.04 (-1.03)	-0.02 (-0.44)
Five-Year Return Loser	0.32 (3.34)	0.09 (1.00)	0.19 (2.28)	0.05 (0.62)	0.08 (1.08)	-0.01 (-0.16)	0.06 (1.01)	-0.01 (-0.12)	0.16 (2.26)	0.03 (0.42)
Five-Year Low Winner	-0.17 (-1.69)	-0.19 (-1.88)	-0.29 (-3.25)	-0.35 (-3.86)	-0.17 (-2.10)	-0.25 (-3.06)	-0.09 (-1.01)	-0.18 (-2.11)	-0.14 (-1.67)	-0.21 (-2.51)
Five-Year Low Loser	-0.05 (-0.77)	-0.06 (-0.82)	-0.01 (-0.21)	-0.02 (-0.34)	0.01 (0.14)	0.01 (0.21)	-0.01 (-0.11)	0.04 (0.67)	-0.02 (-0.38)	0.00 (0.01)

Panel B

$$R_{it} = b_{0jt} + b_{1jt} R_{i,t-1} + b_{2jt} size_{i,t-1} + b_{3jt} 52wkhW_{i,t-j} + b_{4jt} 52wkhL_{i,t-j} + b_{5jt} TRW_{i,t-j} + b_{6jt} TRL_{i,t-j} + b_{7jt} EWGOW_{i,t-j} + b_{8jt} EWGOL_{i,t-j} + e_{ijt}$$

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Five-Year Return Winner	-0.11 (-2.44)	-0.06 (-1.33)	-0.02 (-0.49)	-0.01 (-0.14)	-0.04 (-0.76)	-0.04 (-0.85)	-0.05 (-1.03)	-0.05 (-1.03)	-0.06 (-1.66)	-0.04 (-1.29)
Five-Year Return Loser	0.33 (3.45)	0.10 (1.13)	0.21 (2.51)	0.07 (0.84)	0.08 (1.08)	-0.01 (-0.15)	0.06 (0.99)	-0.01 (-0.13)	0.16 (2.38)	0.04 (0.54)
Equal-Weighted Gain Winner	-0.13 (-1.39)	-0.14 (-1.46)	-0.28 (-3.56)	-0.32 (-4.09)	-0.15 (-2.11)	-0.22 (-3.05)	-0.08 (-1.06)	-0.16 (-2.19)	-0.12 (-1.74)	-0.18 (-2.55)
Equal-Weighted Gain Loser	-0.06 (-0.93)	-0.06 (-0.87)	-0.05 (-0.74)	-0.05 (-0.79)	0.01 (0.21)	0.02 (0.29)	-0.01 (-0.22)	0.03 (0.53)	-0.03 (-0.60)	-0.01 (-0.17)

Table VI - Equal Weighted Gain/Loss and Measures of Embedded Gains

Each month between January 1968 and December 2001, 60 ($j=1,...,60$) cross-sectional regressions are estimated where R_{it} is the return to stock i in month t, $R_{i,t-1}$ and $size_{i,t-1}$ are the return and the natural logarithm of market capitalization of stock i in month $t-1$ net of the month $t-1$ cross-sectional mean; $52wkhW_{i,t-j}$ ($52wkhL_{i,t-j}$) is the 52-week high winner (loser) dummy that takes the value of 1 if the 52-week high measure for stock i is ranked in the top (bottom) 30% in month $t-j$, and zero otherwise. The 52-week high measure in month $t-j$ is the ratio of price level in month $t-j$ to the maximum price achieved in months $t-j-12$ to $t-j$. Similar to the construction of the 52-week high winner and loser dummies, the measures $EWGLW$ and $EWGLL$ are winner and loser dummies defined with respect to $EWGL$ using a 30% cutoff. $EWGL_{i,t-j}$ is the gain net of losses embedded in stock i on date $t-j$ under the assumption that the shares are acquired uniformly over the period from $t-j-60$ to $t-j$. The measures $FYLW$ and $FYLL$ in Panel A ($EWGOW$ and $EWGOL$ in Panel B) are winner and loser dummies defined with respect to FYL ($EWGO$) using a 30% cutoff as in Table V. The coefficient estimates of a given independent variable are averaged over $j=1,...,12$ for columns labeled (1,12), $j=13,...,25$, for columns labeled (13,24),..., $j=1,...,60$ for columns labeled (1,60). The numbers reported in the table are the time-series averages of these averages. They are in percent per month. The accompanying t-statistics are calculated from the time series. Coefficients on control variables are omitted for brevity.

Panel A

$$R_{it} = b_{0jt} + b_{1jt} R_{i,t-1} + b_{2jt} size_{i,t-1} + b_{3jt} 52wkhW_{i,t-j} + b_{4jt} 52wkhL_{i,t-j} + b_{5jt} EWGLW_{i,t-j} + b_{6jt} EWGLL_{i,t-j} + b_{7jt} FYLW_{i,t-j} + b_{8jt} FYLL_{i,t-j} + e_{ijt}$$

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Equal-Weighted Gain/Loss Winner	0.01	0.12	-0.10	-0.05	-0.10	-0.10	-0.07	-0.08	-0.05	-0.02
	(0.15)	(2.16)	(-1.84)	(-0.88)	(-2.11)	(-1.97)	(-1.57)	(-1.82)	(-1.42)	(-0.59)
Equal-Weighted Gain/Loss Loser	0.27	-0.01	0.16	-0.04	0.07	-0.06	0.08	0.00	0.14	-0.02
	(2.58)	(-0.12)	(1.62)	(-0.48)	(0.87)	(-0.68)	(1.11)	(-0.02)	(1.71)	(-0.32)
Five-Year Low Winner	-0.21	-0.31	-0.24	-0.33	-0.14	-0.22	-0.09	-0.16	-0.13	-0.22
	(-2.25)	(-3.23)	(-2.64)	(-3.59)	(-1.71)	(-2.68)	(-1.10)	(-1.99)	(-1.62)	(-2.71)
Five-Year Low Loser	-0.07	-0.02	-0.02	0.01	0.01	0.06	0.03	0.09	-0.03	0.03
	(-0.98)	(-0.25)	(-0.33)	(0.17)	(0.20)	(0.83)	(0.39)	(1.24)	(-0.47)	(0.39)

Panel B

$$R_{it} = b_{0jt} + b_{1jt} R_{i,t-1} + b_{2jt} size_{i,t-1} + b_{3jt} 52wkhW_{i,t-j} + b_{4jt} 52wkhL_{i,t-j} + b_{5jt} EWGLW_{i,t-j} + b_{6jt} EWGLL_{i,t-j} + b_{7jt} EWGOW_{i,t-j} + b_{8jt} EWGOL_{i,t-j} + e_{ijt}$$

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Equal-Weighted Gain/Loss Winner	0.06	0.18	-0.06	0.00	-0.06	-0.05	-0.05	-0.05	-0.02	0.02
	(1.05)	(3.21)	(-0.98)	(0.02)	(-1.25)	(-0.98)	(-1.14)	(-1.10)	(-0.52)	(0.48)
Equal-Weighted Gain/Loss Loser	0.27	-0.01	0.18	-0.02	0.07	-0.05	0.08	0.01	0.15	-0.01
	(2.68)	(-0.06)	(1.92)	(-0.21)	(0.92)	(-0.61)	(1.23)	(0.12)	(1.94)	(-0.13)
Equal-Weighted Gain Winner	-0.25	-0.34	-0.28	-0.37	-0.17	-0.25	-0.10	-0.18	-0.16	-0.25
	(-2.91)	(-3.93)	(-3.39)	(-4.40)	(-2.30)	(-3.40)	(-1.36)	(-2.42)	(-2.12)	(-3.29)
Equal-Weighted Gain Loser	-0.07	-0.01	-0.05	-0.01	0.02	0.07	0.02	0.08	-0.04	0.02
	(-0.95)	(-0.16)	(-0.68)	(-0.13)	(0.34)	(0.92)	(0.33)	(1.16)	(-0.61)	(0.25)

Table VII – Traditional Measure and Measures of Embedded Gains, Risk-Adjusted Returns

Each month between January 1968 and December 2001, 60 (j=1,...,60) cross-sectional regressions are estimated where R_{it} is the return to stock i in month t, $R_{i,t-1}$ and $size_{i,t-1}$ are the return and the natural logarithm of market capitalization of stock i in month $t-1$ net of the month $t-1$ cross-sectional mean; $52wkhW_{i,t-j}$ ($52wkhL_{i,t-j}$) is the 52-week high winner (loser) dummy that takes the value of 1 if the 52-week high measure for stock i is ranked in the top (bottom) 30% in month $t-j$, and zero otherwise. The 52-week high measure in month $t-j$ is the ratio of price level in month $t-j$ to the maximum price achieved in months $t-j$-12 to $t-j$. Similar to the construction of the 52-week high winner and loser dummies, the measures TRW and TRL are winner and loser dummies defined with respect to TR using a 30% cutoff. $TR_{i,t-j}$ is the five-year return for stock i over the period from $t-j$-60 to $t-j$. The measures $FYLW$ and $FYLL$ in Panel A ($EWGOW$ and $EWGOL$ in Panel B) are winner and loser dummies defined with respect to FYL ($EWGO$) using a 30% cutoff as in Table V. The coefficient estimates of a given independent variable are averaged over j=1,...,12 for columns labeled (1,12), j=13,...,25, for columns labeled (13,24),..., j=1,...,60 for columns labeled (1,60). To obtain risk-adjusted returns, we further run times series regressions of these averages (one for each average) on the contemporaneous Fama-French factor realizations to hedge out the factor exposure. The numbers reported for risk adjusted returns are intercepts from these time-series regressions. They are in percent per month and their t-statistics are in parentheses. Coefficients on control variables are omitted for brevity.

Panel A

$$R_{it} = b_{0jt} + b_{1jt}R_{i,t-1} + b_{2jt}size_{i,t-1} + b_{3jt}52wkhW_{i,t-j} + b_{4jt}52wkhL_{i,t-j} + b_{5jt}TRW_{i,t-j} + b_{6jt}TRL_{i,t-j} + b_{7jt}FYLW_{i,t-j} + b_{8jt}FYLL_{i,t-j} + e_{ijt}$$

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Five-Year Return Winner	-0.01 (-0.19)	0.01 (0.15)	0.01 (0.28)	0.03 (0.58)	-0.02 (-0.35)	-0.01 (-0.27)	-0.04 (-0.71)	-0.04 (-0.78)	-0.01 (-0.33)	-0.01 (-0.18)
Five-Year Return Loser	0.19 (2.32)	0.05 (0.66)	0.14 (1.94)	0.06 (0.84)	0.04 (0.59)	-0.01 (-0.20)	0.04 (0.63)	-0.01 (-0.09)	0.08 (1.56)	0.01 (0.19)
Five-Year Low Winner	-0.12 (-2.17)	-0.15 (-2.73)	-0.28 (-5.15)	-0.33 (-6.22)	-0.20 (-3.74)	-0.24 (-4.66)	-0.11 (-2.05)	-0.17 (-3.08)	-0.14 (-3.51)	-0.19 (-4.94)
Five-Year Low Loser	-0.01 (-0.11)	-0.03 (-0.75)	-0.01 (-0.32)	-0.03 (-0.75)	0.00 (0.06)	0.00 (0.11)	0.04 (0.98)	0.06 (1.24)	0.01 (0.23)	0.01 (0.35)

Panel B

$$R_{it} = b_{0jt} + b_{1jt}R_{i,t-1} + b_{2jt}size_{i,t-1} + b_{3jt}52wkhW_{i,t-j} + b_{4jt}52wkhL_{i,t-j} + b_{5jt}TRW_{i,t-j} + b_{6jt}TRL_{i,t-j} + b_{7jt}EWGOW_{i,t-j} + b_{8jt}EWGOL_{i,t-j} + e_{ijt}$$

Five-Year Return Winner	-0.05 (-1.10)	-0.04 (-0.83)	-0.01 (-0.13)	0.00 (0.01)	-0.04 (-0.81)	-0.04 (-0.86)	-0.05 (-1.08)	-0.06 (-1.23)	-0.03 (-1.04)	-0.04 (-1.06)
Five-Year Return Loser	0.20 (2.44)	0.06 (0.80)	0.16 (2.20)	0.08 (1.10)	0.04 (0.64)	-0.01 (-0.15)	0.04 (0.66)	0.00 (-0.06)	0.09 (1.74)	0.02 (0.37)
Equal-Weighted Gain Winner	-0.07 (-1.30)	-0.09 (-1.55)	-0.27 (-5.20)	-0.31 (-5.96)	-0.17 (-3.29)	-0.22 (-4.11)	-0.09 (-1.70)	-0.14 (-2.70)	-0.12 (-3.26)	-0.16 (-4.49)
Equal-Weighted Gain Loser	-0.01 (-0.29)	-0.03 (-0.74)	-0.05 (-1.19)	-0.07 (-1.61)	0.00 (0.10)	0.00 (0.10)	0.03 (0.79)	0.04 (0.98)	-0.01 (-0.26)	0.00 (-0.10)

Table VIII – Traditional Measure and Its Interactions with Five-Year High and Five-Year Low, Hong Kong Data

Each month between January 1985 and December 2000, 60 (j=1,…,60) cross-sectional regressions of the following form are estimated:

$$R_{it} = b_{0jt} + b_{1jt} R_{i,t-1} + b_{2jt} size_{i,t-1} + b_{3jt} TRW_{i,t-j} + b_{4jt} TRL_{i,t-j} + b_{5jt} 52wkhW_{i,t-j} + b_{6jt} 52wkhL_{i,t-j} + b_{7jt} TRW_{i,t-j}*FYHW_{i,t-j} + b_{8jt} TRL_{i,t-j}*FYLL_{i,t-j} + e_{ijt}$$

where R_{it} is the return to stock i in month t, $R_{i,t-1}$ and $size_{i,t-1}$ are the return and the natural logarithm of market capitalization of stock i in month $t-1$ net of the month $t-1$ cross-sectional mean; $52wkhW_{i,t-j}$ ($52wkhL_{i,t-j}$) is the 52-week high winner (loser) dummy that takes the value of 1 if the 52-week high measure for stock i is ranked in the top (bottom) 30% in month $t-j$, and zero otherwise. The 52-week high measure in month $t-j$ is the ratio of price level in month $t-j$ to the maximum price achieved in months $t-j-12$ to $t-j$. Similar to the construction of the 52-week high winner and loser dummies, the measures . FYLL is the loser dummy defined with respect to FYL using a 30% cutoff. $FYL_{i,t-j}$ is the price of stock i at time $t-j$ relative to its five year low over the period from $t-j-60$ to $t-j$. FYHW is the winner dummy defined with respect to FYH using a 30% cutoff. $FYH_{i,t-j}$ is the price of stock i at time $t-j$ relative to its five year high over the period from $t-j-60$ to $t-j$. The measures TRW and TRL are winner and loser dummies defined with respect to TR using a 30% cutoff. $TR_{i,t-j}$ is the five-year return for stock i over the period from $t-j-60$ to $t-j$. The coefficient estimates of a given independent variable are averaged over j=1,…,12 for columns labeled (1,12), j=13,…,24, for columns labeled (13,24),…, j=1,…,60 for columns labeled (1,60). The numbers reported in the table are the time-series averages of these averages. They are in percent per month. The accompanying t-statistics are calculated from the time series.

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Intercept	1.91	2.18	1.80	2.15	1.31	1.76	0.98	1.57	0.75	1.39
	(2.24)	(2.49)	(1.89)	(2.23)	(1.24)	(1.67)	(0.81)	(1.30)	(0.55)	(1.03)
$R_{i,t-1}$	3.15	3.81	3.16	3.64	3.19	3.66	0.20	0.08	0.64	0.83
	(1.53)	(1.73)	(1.53)	(1.71)	(1.50)	(1.70)	(0.11)	(0.04)	(0.29)	(0.37)
Size	-0.25	-0.30	-0.48	-0.49	-0.47	-0.45	-0.51	-0.50	-0.58	-0.61
	(-1.47)	(-1.59)	(-2.18)	(-2.06)	(-1.93)	(-1.71)	(-1.73)	(-1.56)	(-2.18)	(-2.10)
Five-Year Return Winner	-0.12	-0.02	0.00	0.12	0.04	0.13	0.52	0.61	0.35	0.47
	(-0.45)	(-0.07)	(-0.01)	(0.42)	(0.13)	(0.41)	(1.29)	(1.42)	(1.19)	(1.57)
Five-Year Return Loser	0.45	0.54	0.15	0.34	0.14	0.33	0.46	0.69	0.36	0.56
	(1.09)	(1.22)	(0.30)	(0.66)	(0.26)	(0.60)	(0.84)	(1.19)	(0.74)	(1.12)
52-Wk High Winner	-0.13	-0.18	-0.39	-0.36	-0.06	-0.13	-0.06	-0.16	-0.15	-0.24
	(-0.54)	(-0.70)	(-1.83)	(-1.60)	(-0.27)	(-0.53)	(-0.21)	(-0.51)	(-0.77)	(-1.17)
52-Wk High Loser	0.30	0.34	0.49	0.53	0.68	0.74	0.26	0.36	0.72	0.80
	(0.96)	(1.03)	(1.26)	(1.28)	(2.01)	(2.03)	(0.75)	(0.98)	(2.00)	(2.06)
Five-Year Return Winner* Five-Year High Winner	-0.11	-0.08	0.75	0.67	0.91	0.85	1.11	1.09	0.86	0.85
	(-0.39)	(-0.29)	(2.23)	(1.91)	(2.21)	(1.99)	(2.68)	(2.46)	(2.49)	(2.36)
Five-Year Return Loser* Five-Year Low Loser	-0.57	-0.55	-0.60	-0.68	-0.52	-0.60	-0.21	-0.32	-0.37	-0.46
	(-1.57)	(-1.42)	(-1.22)	(-1.27)	(-0.87)	(-0.93)	(-0.36)	(-0.50)	(-0.99)	(-1.16)

Table IX – Traditional Measure and Measures of Embedded Gains, Hong Kong Data

Each month between January 1985 and December 2000, 60 ($j=1,…,60$) cross-sectional regressions are estimated where R_{it} is the return to stock i in month t, $R_{i,t-1}$ and $size_{i,t-1}$ are the return and the natural logarithm of market capitalization of stock i in month $t-1$ net of the month $t-1$ cross-sectional mean; $52wkhW_{i,t-j}$ ($52wkhL_{i,t-j}$) is the 52-week high winner (loser) dummy that takes the value of 1 if the 52-week high measure for stock i is ranked in the top (bottom) 30% in month $t-j$, and zero otherwise. The 52-week high measure in month $t-j$ is the ratio of price level in month $t-j$ to the maximum price achieved in months $t-j-12$ to $t-j$. Similar to the construction of the 52-week high winner and loser dummies, the measures TRW and TRL are winner and loser dummies defined with respect to TR using a 30% cutoff. $TR_{i,t-j}$ is the five-year return for stock i over the period from $t-j-60$ to $t-j$. The measures $FYLW$ and $FYLL$ in Panel A ($EWGOW$ and $EWGOL$ in Panel B) are winner and loser dummies defined with respect to FYL ($EWGO$) using a 30% cutoff. $FYL_{i,t-j}$ is the price of stock i at time $t-j$ relative to its five year low over the period from $t-j-60$ to $t-j$. $EWGO_{i,t-j}$ is the gain only, excluding losses, embedded in stock i on date $t-j$ under the assumption that the shares are acquired uniformly over the period from $t-j-60$ to $t-j$. The coefficient estimates of a given independent variable are averaged over $j=1,…,12$ for columns labeled (1,12), $j=13,…,25$, for columns labeled (13,24),…, $j=1,…,60$ for columns labeled (1,60). The numbers reported in the table are the time-series averages of these averages. They are in percent per month. The accompanying t-statistics are calculated from the time series. Coefficients on control variables are omitted for brevity.

Panel A

$$R_{it} = b_{0jt} + b_{1jt}R_{i,t-1} + b_{2jt}\,size_{i,t-1} + b_{3jt}\,52wkhW_{i,t-j} + b_{4jt}\,52wkhL_{i,t-j} + b_{5jt}\,TRW_{i,t-j} + b_{6jt}\,TRL_{i,t-j} + b_{7jt}\,FYLW_{i,t-j} + b_{8jt}\,FYLL_{i,t-j} + e_{ijt}$$

	Monthly return (1,12)	Monthly return (1,12) Jan. excluded	Monthly return (13,24)	Monthly return (13,24) Jan. excluded	Monthly return (25,36)	Monthly return (25,36) Jan. excluded	Monthly return (37,48)	Monthly return (37,48) Jan. excluded	Monthly return (1,60)	Monthly return (1,60) Jan. excluded
Five-Year Return Winner	-0.13 (-0.57)	0.02 (0.09)	0.12 (0.44)	0.12 (0.44)	0.45 (1.19)	0.09 (0.43)	0.50 (1.93)	0.44 (1.70)	0.29 (1.37)	0.17 (0.87)
Five-Year Return Loser	0.25 (0.95)	0.36 (1.33)	0.28 (0.96)	0.37 (1.32)	0.05 (0.15)	-0.05 (-0.17)	0.92 (0.93)	0.09 (0.25)	0.60 (1.11)	0.25 (0.87)
Five-Year Low Winner	-0.23 (-0.86)	-0.21 (-0.76)	-0.04 (-0.15)	-0.04 (-0.13)	0.23 (0.65)	0.10 (0.34)	0.42 (1.17)	0.54 (1.50)	0.27 (0.97)	0.33 (1.20)
Five-Year Low Loser	-0.09 (-0.32)	-0.02 (-0.06)	-0.07 (-0.27)	-0.18 (-0.82)	-0.22 (-0.76)	-0.14 (-0.48)	-0.26 (-0.48)	0.20 (0.63)	-0.25 (-0.90)	-0.03 (-0.17)

Panel B

$$R_{it} = b_{0jt} + b_{1jt}R_{i,t-1} + b_{2jt}\,size_{i,t-1} + b_{3jt}\,52wkhW_{i,t-j} + b_{4jt}\,52wkhL_{i,t-j} + b_{5jt}\,TRW_{i,t-j} + b_{6jt}\,TRL_{i,t-j} + b_{7jt}\,EWGOW_{i,t-j} + b_{8jt}\,EWGOL_{i,t-j} + e_{ijt}$$

Five-Year Return Winner	-0.17 (-0.75)	-0.07 (-0.32)	-0.01 (-0.04)	-0.02 (-0.06)	0.49 (1.26)	0.08 (0.35)	0.56 (2.10)	0.52 (1.88)	0.28 (1.31)	0.19 (0.91)
Five-Year Return Loser	0.28 (1.09)	0.38 (1.44)	0.22 (0.79)	0.33 (1.19)	0.09 (0.27)	0.00 (-0.01)	1.00 (1.06)	0.22 (0.63)	0.61 (1.17)	0.28 (1.01)
Equal-Weighted Gain Winner	-0.16 (-0.62)	-0.08 (-0.30)	0.19 (0.69)	0.21 (0.75)	0.18 (0.56)	0.08 (0.31)	0.28 (1.00)	0.32 (1.06)	0.28 (1.27)	0.29 (1.31)
Equal-Weighted Gain Loser	-0.14 (-0.56)	-0.05 (-0.20)	0.10 (0.43)	-0.03 (-0.14)	-0.29 (-1.09)	-0.24 (-0.87)	-0.37 (-0.87)	-0.11 (-0.34)	-0.24 (-1.00)	-0.09 (-0.51)

Footnotes

1. We have also constructed the measure with the reference price (cost basis) described in Grinblatt and Han (2005). But that measure also has very little power predicting winner reversals.

2. Losses accrued in earlier years will affect year-end tax planning because investors do not realize all short-term losses in the years when they occur—the disposition effect. Using individual accounts from a discount brokerage, Odean (1998) estimates that when a trade occurs in an investor account, only 10% of the losses are realized on average, while the average is 15% of gains. This suggests that individual portfolios in his sample contain 90 cents of unrealized losses for every 85 cents of unrealized gains at any point in time. Thus, as time passes and individuals accumulate unrealized gains, they tend to accumulate even more unrealized losses. This behavior need not be irrational. Even a tax optimizer may let losses become long-term because of the way tax laws limit the ability to recognize losses. Losses can be utilized to offset current or future gains, but not past gains. Therefore, an investor who faces transaction costs will not realize a loss for tax reasons unless he/she also has enough *current* gain that the tax savings exceed the transaction cost of selling the loser.

3. The wash sale rule reinforces this behavior. The rule is that after realizing a loss on a security sale, the loss will be disallowed for tax purposes if a substantially identical security is purchased within 30 days. If investors choose stocks deliberately, then the losers they sell are prime candidates for reinvestment. Since portfolio decisions are complementary, the *requirement* imposed by the wash sale rule to delay decisions on reinvesting in losers

previously sold may increase the attractiveness of delaying reinvestment decisions on other stocks until that time also.

4. Estimates of control variables are not reported to minimize clutter in the tables.

5. Using the 33 January observations, the January return to *FYL* winners is 0.54% significant at the 10% level ($t = 1.70$); the January return to *EWGO* winners is 0.43% ($t = 1.50$).

6. We examined this two different ways. First, Panel B of Table 1 in DeBondt and Thaler (1987) clearly shows that long term reversals mainly concentrate in years 1 through 3, and the reversals in years 4 and 5 are negligible. We estimated three-year reversals, those associated with (1,36) in our notation, and repeated the analyses that correspond to our Tables V and VI. All the tax-motivated measures have t-statistics greater than two with or without January returns. For example, at the (1,36) horizon, the Equal Weighted Gain Only measure of Table VI has a t-statistic of 2.24 (with January) and 2.75 (without January). Second, most of the literature uses less inclusive cutoffs in identifying winners and losers than we do. We use 30% cutoffs, whereas a 10% cutoff is the least extreme used in the literature, and most studies use a 5% cutoff or the top and bottom 50 stocks. Repeating the analyses in these tables using a 10% cutoff, all tax motivated measures have t-statistics above two in the (1,60) window with or without January returns.

7. We are grateful to Ken French for providing the Fama-French factors.

8. Although the risk-adjusted returns we report allow for factor loadings that vary relative to the portfolio formation month, they do not allow for variation in calendar time. We also used a dynamic risk adjustment procedure similar to that of Grundy and Martin (2001) and the results were qualitatively similar.

9. The intercepts in the risk-adjusted regressions are uniformly insignificant, as they should be if the Fama-French factors explain returns to neutral portfolios of Hong Kong stocks.

10. We are grateful to the referee for suggesting this test.

References

Ball, Ray, and S.P. Kothari, 1989, Nonstationary Expected Returns: Implications for Tests of Market Efficiency and Serial Correlation in Returns, *Journal of Financial Economics,* 25, 51-74.

Ball, Ray, S.P. Kothari and Jay Shanken, 1995, Problems in Measuring Portfolio Performance: An Application to Contrarian Investment Strategies, *Journal of Financial Economics,* 38, 79-107.

Barberis, Nicholas, Andrei Shleifer and Robert Vishny, 1998, "A Model of Investor Sentiment*," Journal of Financial Economics*, 49, 307-343.

Berk, Jonathan, Richard Green and Vasant Naik, 1999, "Optimal Investment, Growth Options and Security Returns," *Journal of Finance,* 54, 1553-1607.

Brav, Alon, and J.B. Heaton, 2002, "Competing Theories of Financial Anomalies," *Review of Financial Studies,* 15, 575-606.

Chan, K.C., 1988, "On the Contrarian Investment Strategy," *Journal of Business,* 61, 147-163.

Chopra, Naveen, Josef Lakonishok and Jay Ritter, 1992, "Measuring Abnormal Performance: Do Stocks Overreact?" *Journal of Financial Economics,* 31, 235-268.

Conrad, Jennifer, Michael Cooper and Gautam Kaul, 2003, "Value Versus Glamor," *Journal of Finance*, 58, 1969-1995.

Constantinides, George, 1984, "Optimal Stock Trading with Personal Taxes: Implications for Prices and the Abnormal January Returns," *Journal of Financial Economics,* 13, 65-89.

Daniel, Kent, David Hirshleifer and A. Subrahmanyam, 1998, "Investors Psychology and Security Market Under- and Over-Reactions," *Journal of Finance*, 53, 1839-1885.

DeBondt, Werner, and Richard Thaler, 1985, "Does the Stock Market Overreact?," *Journal of Finance*, 40, 793-805.

DeBondt, Werner, and Richard Thaler, 1987, "Further Evidence On Investor Overreaction and Stock Market Seasonaility," *Journal of Finance*, 42, 557-581.

D'Mello, Ranjan, Steven Ferris and Chuan-Yang Hwang, 2003, "The Tax-Loss Selling Hypothesis, Market Liquidity, and Price Pressure around the Turn-of-the-Year," *Journal of Financial Markets*, 6, 73-98.

Fama, Eugene, 1976, *Foundations of Finance: Portfolio Decisions and Securities Prices,* Basic Books Inc, New York.

Fama, Eugene, and James MacBeth, 1973, "Risk, Return and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607-636.

Fama, Eugene, and Kenneth French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics,* 33, 3-53.

Fama, Eugene, and Kenneth French, 1996, "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance,* 51, 55-84.

George, Thomas, and Chuan-Yang Hwang, 2004, "The 52-Week High and Momentum Investing," *Journal of Finance*, 59, 2145-2176.

Griffiths, Mark, and Robert White, 1993, "Tax Induced Trading and the Turn-of-the-year Anomaly: An Intraday Study," *Journal of Finance*, 48, 575-598.

Grinblatt, Mark, and Bing Han, 2005, "Prospect Theory, Mental Accounting and Momentum," *Journal of Financial Economics*, 78, 311-339.
.
Grinblatt, Mark, and Matti Keloharju, 2004, "Tax Loss Trading and Wash Sales," *Journal of Financial Economics,* 71, 51-76.

Grinblatt, Mark, and Tobias Moskowitz, 2004, "Predicting Stock Price Movements from Past Returns: the Role of Consistency and Tax Loss Selling," *Journal of Financial Economics*, 71, 541-579.

Grundy, Bruce, and Spencer Martin, 2001, "Understanding the Nature of the Risks and the Source of the Rewards to Momentum Investing," *Review of Financial Studies,* 14, 29-78.

Hong, Harrison, and Jeremy Stein, 1999, "A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets," *Journal of Finance*, 54, 2143-2184.

Ibbotson, Roger, 1975, "Price Performance of Common Stock New Issues," *Journal of Financial Economics,* 2, 235-272.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, "Returns to Buying Winners and Selling Losers: Implications for Market Efficiency," *Journal of Finance*, 48, 65-91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations," *Journal of Finance*, 56, 699-718.

Klein, Peter, 1999, "The Capital Gains Lock-In Effect and Equilibrium Returns," *Journal of Public Economics,* 71, 355-378.

Klein, Peter, 2001, "The Capital Gains lock-In Effect and Long-Horizon Return Reversal," *Journal of Financial Economics,* 59, 33-62.

Lewellen, Jonathan, and Jay Shanken, 2002, "Learning, Asset-Pricing Tests, and Market Efficiency," *Journal of Finance,* 57, 1113-1145.

Moskowitz, Tobias, and Mark Grinblatt, 1999, "Do Industries Explain Momentum?," *Journal of Finance,* 54, 1249-1290.

Odean, Terrence, 1998, "Are Investors Reluctant to Realize their Losses?" *Journal of Finance,* 53, 1775-1798.

Ritter, Jay, 1988, "The Buying and Selling Behavior of Individual Investors at the Turn of the Year," *Journal of Finance,* 43, 701-717.

Roll, Richard, 1983, "Vas ist das? The Turn-of-the-Year Effect and the Return Premium of Small Firms," *Journal of Portfolio Management* 9, 18-28.

Schultz, Paul, 1985, "Personal Income Taxes and the January Effect: Small Firm Stock Returns Before the War Revenue Act of 1917: A Note," *Journal of Finance,* 40, 333-344.