This is a repository copy of *Capital Gains Overhang with a Dynamic Reference Point*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/146448/

Version: Accepted Version

Capital Gains Overhang with a Dynamic Reference Point

Christopher Riley[a], Barbara Summers[b], Darren Duxbury[c]

Abstract

Financial models incorporating a reference point, such as the Capital Gains Overhang (CGO) model, typically assume it is fixed at the purchase price. Combining experimental and market data, this paper examines whether such models can be improved by incorporating reference point adjustment. Using real stock prices over horizons from 6-months to 5-years, experimental evidence demonstrates that a number of salient points in the prior share price path are key determinants of the reference point, in addition to the purchase price. Market data testing is then undertaken using the CGO model. We show that composite CGO variables, created using a mix of salient points with weights determined in the experiment, have greater predictive power than the traditional CGO variable in both cross-sectional US equity return analysis and when analyzing the performance of double-sorted portfolios. In addition, future trading volume is more sensitive to changes in the composite CGO variables than to the traditional CGO, further emphasizing the importance of adjusting reference points.

[a] University of Leicester School of Business, email: C.Riley@leicester.ac.uk
[b] Centre for Decision Research, Leeds University Business School, email: bs@lubs.leeds.ac.uk
[c] Newcastle University Business School, email: Darren.Duxbury@newcastle.ac.uk

1. Introduction

The reference point is a central feature of Prospect Theory (Kahneman and Tversky, 1979), but it has received less attention than other aspects of the theory, such as the differing shape of the value function when in a position of gain or loss. The choice of reference point has direct implications for the determination of gains and losses, hence it impacts all kinds of financial behavior ranging from firm-based capital investment (Whyte, 1986) and strategic decisions (Bamberger and Fiegenbaum, 1996), to stock investor behavior (Shefrin and Statman, 1985). A key motivation for our study is to place the spotlight back on the role of the reference point, to show that it adjusts as investors experience movements in the stock price,[1] and to demonstrate how incorporating this adjustment can improve financial models. To achieve this we combine experimental and market data methods, by identifying how reference points adapt in an experiment and then using the derived parameters to test the predictive power of the model in market data. We demonstrate that treating the reference point of an investor as something that changes over time, rather than remaining fixed at the purchase price, can improve the predictive power of the Capital Gains Overhang (CGO) model developed by Grinblatt and Han (2005).

The first aim of our study is to explore the role of alternative salient prices, in addition to the purchase price, in the formation of the reference point. Prior literature, such as work on the disposition effect (Shefrin and Statman, 1985), suggests that investors adopt the purchase price as the reference point, which is then fixed over the life of the investment. A purchase-price based reference point is consistent with the tendency of investors to exhibit a status-quo bias, which is prevalent across a wide range of decisions (Samuelson and Zeckhauser, 1988). A number of field data based studies also provide support for a purchase-price based reference point such as Kaustia (2010), which shows the investor's propensity to sell a stock jumps at the

---

[1] Investor behavior is an ideal setting in which to investigate reference point adjustment due to the higher frequency with which stock markets feed back information (stock price changes) on prior decisions (trading decisions) relative to other decision contexts such as capital budgeting or strategic investment.

purchase price but is constant or only slightly increasing across a range of gains and losses relative to the purchase price, and Ben-David and Hirshleifer (2012), which shows that the propensity to sell follows a V-shaped pattern around the purchase price.

In the context of stock trading, the issue in assuming a static, purchase price based reference point, is that stocks can be held for a prolonged period of time. As the stock price moves, the question then becomes: will the investor update their reference point in response to the new price information, or will the reference point remain fixed at the purchase price? Papers from Chen and Rao (2002) and Arkes, Hirshleifer, Jiang and Lim (2008) suggest that investors do indeed update their reference point as new price information arrives. Arkes et al. (2008) show that just over half of the gain is adjusted for in the new reference point and slightly less for losses.

A number of different salient points could play a role in the reference point updating process. Both Heyman, Mellers, Tishcenko and Schwartz (2004) and Gneezy (2005), using experimental methods, find that historic highs are key determinants of the reference point and there is further evidence for the importance of the high price within studies using market data. Kaustia (2004) suggests that maximums and minimums, attained over the prior month, are key price points for investors in new IPOs, with stock price moves to these points triggering higher trading activity. Both Heath, Huddart and Lang (1999) and Poteshman and Serbin (2003) find that exercise of stock options greatly increases when the stock price exceeds the maximum over the year. There is also support for the 52-week high in addition to the historic high. Huddart, Lang and Yetman (2009) find volume increases for stock prices close to the 52-week high, while George and Hwang (2004) find that stocks near to their 52-week high, in percentage terms, tend to be underpriced and subsequently outperform, relative to stocks that are far from the 52-week high. The authors suggest that at price levels close to the 52-week high, traders are reluctant to bid the stock price higher, as it is a key

reference point in the minds of investors. This work on different reference points indicates that studies exploring how multiple salient prices influence reference point formation would be informative.

Baucells, Weber and Welfens (2011) measures the impact of several of these salient points on the reference point and finds that it is determined by a mix of the purchase, maximum, minimum, average and final prices. In our study, we adapt the Baucells et al. (2011) framework by using stock price charts created from real market data which are up to 5 years in length, giving a greater time range than those in Baucells et al. (2011). The results of our experimental study suggest that intermediate highs and lows are important in reference point formation, in addition to the purchase and final prices, in line with the findings in Baucells et al. (2011). In addition, we show the impact of 52-week highs and lows in the determination of investor reference points when longer time periods are considered. This is a new result that is only observed by using long-term charts, whereas previous experimental studies using shorter time periods have not identified the effect. This is not only of conceptual importance, but is also important practically, given the use of longer-term price charts in the real world.

The second aim of our paper is to demonstrate the effect of alternative reference points within a market data model, which uses reference prices to predict future stock price returns. If investor reference points are fixed at the purchase price then it would be surprising if some of the alternative salient points, discussed above, do also have predictive power for future returns. The specific model that we adopt is the CGO model developed by Grinblatt and Han (2005), who show that CGO is a key variable that generates the underlying profitability of a momentum trading strategy (Jegadeesh and Titman, 1993). Stocks which have positive CGO (final price above reference point) tend to be under-priced and subsequently outperform, whereas stocks that have negative CGO (final price below reference point) tend to be over-priced and subsequently underperform. A purchase price based reference point is assumed, while the probability that a stock is bought on a particular day, establishing a new purchase price, is approximated using a stock's turnover

ratio. As such, the Grinblatt and Han (2005) CGO variable takes into account external reference point updating, which occurs due to the sale of the stock and purchase by a new stockholder. Under conditions of zero turnover, however, no updating occurs at all in their model, as the stock is not changing hands between investors.

Our contention is that internal reference point updating on the part of existing stockholders also occurs, even if no stocks are traded, responding to developments in the stock price over time. Internal reference point updating cannot be directly detected in market data however and so our experimental framework is vital in identifying important salient prices. Our experimental results suggest that prices such as the maximum, minimum and 52-week variables are important determinants of the reference point, in addition to the purchase price. We then apply the concept to market data by calculating alternative CGO variables using the salient points identified in the experiment (the maximum, minimum, average, and 52-week variables) and comparing them to the standard model. We show that these alternative CGO variables are just as predictive of 1 month ahead returns as the traditional CGO variable, calculated using the purchase price.

The third aim of our paper is to investigate if CGO variables that use more accurate reference points, formed using a mix of salient points, are better predictors of returns than the traditional CGO that uses the purchase price alone. This would demonstrate the reality of reference point adjustment in real-world stock price data and confirm the earlier experimental results. We use the coefficients from the experiment to create composite-CGO variables, formed by weighting a number of salient points in a price chart to create a more accurate reference point. Our first composite variable, CGOCom1, uses the purchase, maximum and minimum prices, while the second variable, CGOCom2, includes the 52-week prices in place of the maximum and minimum. In cross-sectional regressions of US equity returns where both the composite-CGO variables and the traditional CGO variable are present, we find that the composite-CGO variables are

better predictors of future returns than the traditional variable, with the traditional CGO variable no longer a positive predictor of returns when either of the CGO-Composite variables are included in the regression. Furthermore, we carry out double sorts of portfolios, mirroring the approach in Grinblatt and Han (2005), by both traditional CGO and our preferred CGO-Composite variable incorporating 52-week prices, CGOCom2. We find that that CGO is rarely predictive of returns after stocks are first sorted by CGOCom2, whereas CGOCom2 is usually predictive of returns even if stocks are first sorted by CGO.

Our fourth aim is to subject our findings to a broader test premised on the implication that if reference points adjust and composite reference points are predictive of returns, then trading volume should also be more sensitive to composite reference points than to the purchase prices.[2] To this end, we complement the returns analysis by carrying out volume tests and find that the composite CGO variables are also stronger predictors of future trading volume, thus extending the implications of our findings. Specifically, we find that future weekly volume increases almost monotonically in line with CGOCom1 or CGOCom2 when stocks are sorted into deciles based on these variables and volume is more sensitive to the composites than to the traditional CGO variable. The results suggest that future volume as well as future returns are more responsive to CGO variables based on a composite reference point than CGO based on a purchase price based reference point.

We also test if retail investors are more sensitive to reference point affects than institutional traders.[3] If retail investors are more likely to be the irrational traders then the predictive power of CGO and the CGO composite variables should be stronger among the more speculative stocks that are more likely to be traded by these retail investors (Han and Kumar, 2013). We find that, in two out of three of the proxies used for speculative stocks, that both the CGO and CGO composites are robust across the different investor

---

[2] We thank an anonymous reviewer for suggesting this additional line of enquiry.
[3] We thank an anonymous reviewer for suggesting this additional line of enquiry.

segments. In the case of the third proxy for speculative stocks, however, we do find that both the CGO and CGO composites have stronger predictive power among high turnover stocks. An explanation for this may lie in the workings of the CGO model itself, where high turnover tends to refresh the aggregate reference point and temporarily raise returns to the CGO variable.

The implication of our results is that investors take multiple points into consideration when forming a reference point, so that reference points do adjust over time, and adjusted CGO variables are therefore a better predictor of future returns and trading volume than the traditional CGO variable. Aside from the purchase price, key determinants of the reference point are intermediate points of interest such as recent highs and lows. The findings have wider significance, as the CGO model is not the only one that utilizes the concept of a fixed reference point and thus applying adjusted reference points may improve other models. Understanding reference point adjustment is also important for many other concepts in finance, such as the disposition effect. The results also have importance more generally in the management literature, where reference point updating is important but difficult to examine experimentally within an ecologically valid setting and also empirically in light of both the limited (relative) frequency and the potentially widely varying economic contexts within which managerial activities such as capital budgeting and strategic investment decisions are undertaken.

2. Experimental Examination of Reference Point Determination

2.1 Data and Method

In the experiment, we examine the impact of features of the stock price path on the reference point adopted by participants. A repeated measure design is adopted with 30 different chart patterns shown to participants. Order effects are controlled for by randomizing the order of presentation of charts across participants. A data survey company is used to collect online responses. All responses are taken from US Citizens who are

resident in the United States. To ensure that no novices participate in the experiment, all participants are required to have some experience (self-reported) of trading in US stocks or mutual funds, even if this is infrequent. The instructions and example chart for the experiment are available in the appendix.

The experimental approach adopted is based on that of Baucells et al. (2011), with some adjustments to accommodate the longer time frame required for the CGO model, which we use in testing reference point adjustment, as suggested by our experiment, in a market context. An equal number of charts are used with the following lengths: 6 months, 1 year, 2 years, 3 years and 5 years. The increase in chart length, relative to other reference point studies such as Baucells et al. (2011), enables detailed analysis of the impact of recent points on the reference point, such as 52-week highs and lows, as well as overall highs and lows.[4] To ensure that participants recognized the change in time frames, all charts of the same length were presented as a group along with an introductory screen to alert them that the timeframe had changed. The timeframe was also clearly visible in the x-axis of the charts. The order of presentation of groups was randomized as well as the presentation of charts within groups, as indicated above, to counteract order effects.

In Baucells et al. (2011), each point in the graph is presented with a 3-4 second lag, reflecting the later emphasis in their paper on measuring the amount of adaptation at each point of the stock price movement, which may also facilitate the participant experiencing time. Our focus is on determining the likely position of the reference point based on prior stock price movements, which then feeds into the CGO calculation. Given that investors and traders look at a previous stock price path as an entire series rather than as a set of

---

[4] Indeed, Baucells et al. (2011) notes the provocative nature of the conclusion reported therein, that historical peaks seem to matter little for reference points (with historic troughs even less), and calls for further exploration in this regard. We duly provide such exploration in the context of longer price sequences, thus allowing an examination of both historic highs/lows and recent (52-week) highs/lows.

lagged points, our study presents the graphs to participants without delay between points and we elicit one reference point per stock price chart.

The stock price charts were formed from real stock price data using random sampling of equities from the dataset of all US equities from 1963-2016, which is used later in the market data testing section. The use of real stock price data differs from previous experimental studies and most closely replicates the stock price path that investors will observe in real markets. It also avoids bias that may come from artificially generated price series and round numbers that may act as an anchor for participants (Bhattacharya, Holden and Jacobsen, 2012). One drawback to using real stock price data is that multicollinearity between the variables is likely to be high, as it is a natural feature of stock price movements. More rigidly designed charts ensure that multicollinearity can be controlled but would reduce realism and could bias the result in a predictive context, such as that used here, if they are perceived as non-random by participants. We, therefore, use approaches to reduce multicollinearity rather than artificial data.

In terms of capturing reference points from participants, this study uses an adjusted form of the question used in Baucells et al. (2011), 'at what selling price would you feel neutral about the sale of the stock, i.e. be neither happy nor unhappy about the sale'. The adjustment allows for the fact that participants may feel positive and negative emotions at the same time (Cacioppo and Berntson, 1994), and therefore asks participants to consider the balance of their feelings. Therefore, the question we adopt is, *"Your task will be to indicate the selling price at which you would feel neutral (i.e. feel neither predominantly positive nor negative) about selling the stock."* In our study participants are asked for their reference point without any actual sale being involved, allowing for a wider perspective to avoid promoting anchoring on the final price.

Previous research has shown that beliefs about future prices can affect the reference point (Grosshans and Zeisberger, 2017, Hoffmann, Henry and Kalogeras, 2013). Baucells et al. (2011) control for this by

8

informing participants that all possible price changes between €+50 and €−50 are equally likely. Our study did not use this approach because beliefs are part of the way investors form reference points in real life trading and stock price patterns that investors experience can influence these beliefs (Barberis, Shleifer and Vishny, 1998, Rabin and Vayanos, 2010). Thus we look to avoid issues with ecological validity in the CGO calculation.

A feasible range of max +25% and min -25% around the range of the stock price chart was adopted as a limit to establish participant understanding of the experiment. 22 participants, who gave an answer outside this range for a third or more of the charts, were excluded from the results. As an additional check, reference points outside the feasible range of +/-25% were also removed on an individual basis, rather than excluding the whole sample of a participant. These reference points may be invalid due to participant error in one specific chart, but not across the whole experiment. This removed an additional 65 reference points. The final data sample, after exclusions, comprised 169 participants (109 male, 60 female) with a broad distribution of ages between 30-65 years (average = 53 years).

As an additional check, the regression models in section 2.2 are replicated for data screened using the Outlier Sum method (Tibshirani and Hastie, 2007), which uses the distribution of reference points alone to remove outliers. Results are shown in the Electronic Companion in Tables A & B and there is no material difference in the final results between screening based on the range of the charts or by the distribution using the Outlier Sum method.

2.2 Results

Table 1 below shows the characteristics of the 30 charts, along with the average, median and standard deviation of reference points from all participants. All charts are constructed by random sampling from the market dataset used in section 3 so that real price sequences are used.

9

[Insert Table 1]

Regression: Reference Point using Price Variables

To understand the salient prices that determine the reference price and to provide unbiased coefficient estimates for use in the market data testing to follow, our approach to modelling is to begin with a broad set of explanatory variables so as to obtain as the highest explanatory power attainable and then to remove variables in such a way that explanatory power is maintained, but multicollinearity is reduced to acceptable levels. It is important to do this as we want reliable coefficients for market data testing. Models A, B and C in Table 2 use the purchase, maximum, minimum, average and final prices as independent variables (IV's), predicting the reference point as the dependent variable (DV), as shown in Equation (1).

$$Reference\ Point = \ \beta_0 + \beta_1 Purchase + \beta_2 Maximum + \beta_3 Minimum + \beta_4 Average + \beta_5 Final + \varepsilon\ (1)$$

We start with the full set of variables in Model A and then remove IV's while observing the R-squared of the reduced model. The Variance Inflation Factor (VIF) of an IV is calculated by regressing the IV against the other IVs in the model. The R-squared from this regression is then used to calculate the VIF score using Equation (2).

$$VIF = \frac{1}{1-R^2} \quad\quad\quad (2)$$

Models D, E, and F of Table 2 replace the maximum and minimum variables with the 52-week maximum and minimum variables, as shown in Equation (3). Model D begins with the full set of variables, which is gradually reduced by removing IVs. The 52-week variables take account of the investment horizon. For example, over the 6-month charts, the 52-week maximum is the maximum over the 6-month horizon

10

observed and does not take account of prices in the prior 6 months. This distinction is maintained in subsequent market data testing in Section 3.

$$Reference\ Point = \ \beta_0 + \beta_1 Purchase + \beta_2 Max52 + \beta_3 Min52 + \beta_4 Average + \beta_5 Final + \varepsilon \qquad (3)$$

Linear least squares regression is used and robust standard errors are clustered by participant, to mirror the approach in Baucells et al. (2011). In this and all subsequent regressions, variables are taken to be significant if they exceed the 5% significance threshold.

[Insert Table 2]

[Insert Table 3]

In Model A, the purchase, maximum and final price are found to be significant, but the average and minimum prices are insignificant. Table 3 suggests the level of collinearity between the variables in Model A is high, with a mean VIF of 78 and with the maximum, minimum, average and final prices having high VIF scores.

In Models B and C we eliminate one of the two insignificant variables. Model B removes the average variable, which has the largest VIF, while retaining the minimum, and all of the remaining variables are then significant. The R-squared of the model is not reduced relative to Model A and the mean VIF falls to 10. Model C removes the minimum variable while retaining the average. All of the remaining variables are significant with no drop in R-squared. The mean VIF score is higher than model B however, with both the maximum and average having high VIF scores, leaving us with a preference for model B.

The results from Models A, B and C suggest that the purchase, maximum and final prices play a role in determining the reference point, along with the minimum or the average price. These results are in line with the findings of Baucells et al. (2011), although they also found that the average price plays a role as there is no collinearity in their constructed data. We find that Models B or C, with minimum or average included, are close substitutes for each other due to the natural multicollinearity that is present in real share price data. As the VIF scores are lower in Model B however, we construct our first composite reference point, RefCom1, using the weights in Model B. This will be used in the subsequent market data testing section.

In Models D, E and F we replace the maximum and minimum variables with the 52-week maximum (max52) or the 52-week minimum (min52) respectively. In Model D, which includes all the variables, the purchase, max52 and min52 variables are significant but the final and average prices are insignificant. The average VIF score is 40 suggesting that we can remove variables from the regression. Model E removes the average price from the regression. The purchase, max52, and min52 variables remain significant but the final price remains insignificant. As the average VIF score is still high at close to 17, in Model F we also remove the final price. All of the remaining variables are significant and there is no reduction in R-squared, while the average VIF score is reduced to 4.

Model F is our preferred model of the three that introduce the 52-week variables. This model has the same R-squared as model D, which includes all the independent variables, and has a far lower average VIF score. Therefore, coefficients from Model F are used to calculate our second composite reference point, RefCom2, which will be used in the subsequent market data testing section.

In the Electronic Companion, we replicate Table 2 and Table 3 using difference variables shown as Table C and Table D. Each variable is calculated as the percentage deviation of the variable from the final price,

rather than the price itself. This reduces multicollinearity between the variables and leads to lower VIF scores. The results are in line with those in Table 2.

3. Market Data Testing of CGO

The CGO (g) in Grinblatt and Han (2005) is defined as the deviation between the final price (P) and reference price (R), divided by the final price as shown in Equation (4). The final price is lagged by one week, relative to the reference point, to avoid market microstructure events such as the bid-ask bounce that lead to short-term reversal in stock prices (Rosenberg and Rudd, 1982, Da, Liu and Schaumburg, 2013).

$$g_{t-1} = \frac{P_{t-2} - R_{t-1}}{P_{t-2}} \tag{4}$$

It is impossible to calculate the reference point for a given investor directly using market data, as holders buy on different days and therefore have different reference points. To overcome this problem it is necessary to calculate the probability that a stock, currently under ownership, was traded on a particular day. Grinblatt and Han (2005) calculate this probability using stock turnover across 260 weeks of data (5 years in total). By way of example, assuming turnover is 5% on week$_{t-2}$ then the purchase price of that week is given a 5% weight in the reference price (R_{t-1}). For week$_{t-3}$, the turnover for that week again reflects its weight, but some of the buyers in week$_{t-3}$ may also sell during the following week$_{t-2}$. To reflect this, if the turnover in week$_{t-3}$ is 5% then its weight will be 5% * (1 - 5%) = 4.75% to reflect that 5% of the purchases are subsequently sold in week$_{t-2}$.

Theoretically, the reference price is calculated using an infinite number of weeks, but in practice Grinblatt and Han (2005) sum 5 years of weekly turnover adjusted purchase prices and adjust by a constant to make the weights sum to 1. This does not lead to much information loss relative to an infinite calculation, as the

weight given to purchase prices beyond 5 years is typically very small given the high level of weekly turnover for most securities in the market.

The actual adjustment we make to the calculation of the CGO in this study is to exchange the price variable used to calculate the reference point. Equation (5) shows the reference point calculation using the purchase price, as in Grinblatt and Han (2005), to form *RefPurchase*. All other elements of the CGO calculation remain identical to that of Grinblatt and Han (2005), other than that we use daily turnover (V) and daily price information rather than weekly and hence 1260 trading days rather than 260 weeks. Daily data allows the reference point to be calculated more accurately than weekly although it is more computationally intensive. This approach has been adopted by more recent papers that use the CGO model, e.g. Wang, Yan and Yu (2017) or An, Wang, Wang and Yu (2019) .

$$RefPurchase_{t-1} = \frac{1}{k} \sum_{n=1}^{1260} \left(V_{t-1-n} \prod_{\tau=1}^{n-1} [1 - V_{t-1-n+t}] \right) Purchase_{t-1-n} \qquad (5)$$

Notes: Turnover (V) = Daily trading volume/Shares Outstanding, k=constant that makes the weights on past prices sum to one, n=number of trading days from 1 to 1260.

In addition to *RefPurchase*, we use five alternative reference points to create: *RefMax, Refmin, RefMax52, RefMin52* and *RefAverage,* to reflect the salient points that we found to be important in reference point formation. Using Equation (4), the reference points are used to calculate six CGO variables: *CGO, CGOMax, CGOMin, CGOMax52, CGOMin52,* and *CGOAverage.* The maximum or minimum price used in *RefMax* or *Refmin,* at a given point in time, is a function of when the investor bought the security. For example, if the investor bought the security 6 months ago then the maximum or minimum used is that over the last 6 months, as this represents the maximum or minimum over the life of their investment. *RefMax52* and *RefMin52* use the 52-week high or 52-week low respectively. These variables are also a function of when the investor bought the security and so a 6-month holder may have a lower 52-week high than an

investor with a 12 month or longer holding period, as well as a higher low. This mirrors the approach used in the experiment.

We also create two composite reference points formed from a mix of salient points, with weights determined by the experiment. The first composite variable, *RefCom1* shown in Equation 6, is created from Model B of Table 2, which suggests that the reference point is represented by the purchase, maximum, minimum and final prices. *RefCom2* is based on Model F of Table 2, which suggests the reference point is represented by the purchase, 52-week maximum, and 52-week minimum prices. The reference points RefCom1 and RefCom2 are then fed into Equation (4) to create the CGO variables, CGOCom1 and CGOCom2.

$$RefCom1_{t-1} = \frac{1}{k}\sum_{n=1}^{1260}\left(V_{t-1-n}\prod_{\tau=1}^{n-1}[1 - V_{t-1-n+t}]\right)(0.33 * Purchase_{t-1-n} + 0.29 * Max_{t-1-n}$$
$$+ 0.14 * Min_{t-1-n} + 0.23 * Final_{t-1-n}) \tag{6}$$

$$RefCom2_{t-1} = \frac{1}{k}\sum_{n=1}^{1260}\left(V_{t-1-n}\prod_{\tau=1}^{n-1}[1 - V_{t-1-n+t}]\right)(0.43 * Purchase_{t-1-n} + 0.45 * Max52_{t-1-n}$$
$$+ 0.11 * Min52_{t-1-n}) \tag{7}$$

3.1 Data and Method

The market data sample is all US common Stocks (Codes 10 &11) from January 1958 until Dec 2016. NYSE, Amex, and NASDAQ firms are included, although NASDAQ firms have their volume cut in half to compensate for double counting of volume (Anderson and Dyl, 2007). Daily data is used to calculate the CGO variable and is then converted to monthly data for the regressions. The monthly data is from Jan 1963 until Dec 2016, as the CGO variable calculation requires 5 years of data. We convert to a monthly basis, rather than weekly as in Grinblatt and Han (2005), because this is a common frequency for asset pricing

tests and a more common holding period and evaluation period for investors. Stocks are ranked by market capitalization every month as a liquidity screen, with stocks in the bottom-decile rank eliminated for that month. This is to remove the impact of illiquid, untradeable stocks, which could bias the results. There is a total of 68 million firm-day cases in the daily data and around 2.2 million in the monthly data.

The following control variables are used, taken from Grinblatt and Han (2005): *Mom-* is the momentum factor defined as the percentage return over the last 12 month excluding the last month, *STR-* is the short-term reversal factor defined as the percentage return last month, *LTR-* is the long-term reversal factor defined as the percentage return over the last 3 years excluding the last year, *AvgTurn-* is the average of daily stock turnover (daily volume/shares outstanding) over last year and *Mrkcap-* is the log of market cap (stock price*shares outstanding) in units of millions. An additional control variable is included, *BM-* is the log of the book to market ratio, with a minimum lag of 6 months from the reporting date.

Calculation of *CGO* and *LTR* require a minimum of 3 years of data (out of a possible of 5 year) and are set to missing otherwise, as are the reference points: *RefPurchase, RefMax, RefMin, RefMax52, RefMin52, RefAverage, RefCom1 and RefCom2*. Prices are adjusted for stock splits when used to calculate the CGO. In the following regressions, the Fama-Macbeth (Fama and MacBeth, 1973) method is utilized, to mirror the original approach in Grinblatt and Han (2005). Standard errors are corrected using the Newey-West method (Newey and West, 1987) with a lag length of 12. Descriptive statistics for all variables are provided in the Electronic Companion in Tables E & F.

Figure 1 shows how the reference point, based on the purchase price or the second composite, compares to the stock price for the firm IBM (ticker IBM), shown for illustrative purposes. IBM has a turnover of around 80% per annum, which is fairly typical for the sample (the average stock turnover is 86% per annum across

the sample). This means that around 80% of the weight of the reference point is provided by the first 3 years of price data, with older data playing less of a role.

Figure 1: IBM Price & Reference Prices: 1993-2013

3.3 Regression of One Month Ahead Returns against CGO Variables

Table 4 shows regression analysis for one month ahead returns as the dependent variable, various specifications of the CGO variable as independent variables, and the control variables.

$$Ret_{t+1} = \beta_0 + \beta_1 CGO_t + \beta_2 MOM_t + \beta_3 STR_t + \beta_4 LTR_t + \beta_5 Avgturn_t + \beta_6 Mrkcap_t + \beta_7 BM + \varepsilon \qquad (8)$$

[Insert Table 4]

17

Model A features the traditional CGO variable, calculated using the purchase price alone. The results show that the CGO based on purchase price is significant at the 5% level, along with some of the control variables. Whilst the dual significance of CGO and MOM is at odds with the findings of Grinblatt and Han (2005), it is consistent with a more recent study (Wang et al., 2017) that uses daily data to calculate the CGO as we do. Models B to F use the alternative CGO variables calculated using the alternative reference points. All of these variables are significant at the 5% level, suggesting that these CGO variables are also predictive of future returns. All of the six models have a similar Average R-squared. The results suggest that the purchase price is not the only point that is relevant in investor reference point determination, as the alternative specifications of the CGO, using alternative reference points, have similar levels of power to predict one month ahead returns.

[Insert Table 5]

Table 5 shows the regression results of models that include both the traditional CGO variable and composite CGO variables in the same model, to assess which variable retains its significance as a positive predictor of returns. Models B & C include the CGO composite variables along with the traditional CGO. In both cases, the composite CGO variables are positive and significant predictors of returns, while the traditional CGO has a negative coefficient. Both the average R-squared and the adjusted R-squared of Models B & C are greater than Model A, reflecting an increase in predictive power that the addition of the composite variables provide. This result suggests that the composite based CGO measures are better predictors of future returns than the traditional CGO variable. CGOCom2 has more explanatory power than CGOCom1, suggesting that the 52-week high and 52-week low have an important influence on reference points. Model D includes both CGO composite variables *CGOCom1* and *CGOCom2*, and in this instance neither of them is significant due to collinearity. We will confirm the result with double sorted portfolio results in the next section.

18

We also conduct a series of tests to check the robustness of our results in the Electronic Companion.[5] Grinblatt and Han (2005) find that CGO becomes a negative predictor of returns in the month of January due to tax loss selling behavior. In Table G we perform subsample-analysis across the January and February-December months to consider seasonality effects. The results show that our composites exhibit the same seasonality in returns as the original CGO variable, with negative coefficients during the month of January due to tax-loss selling.[6] In Table H, we consider three additional control variables (also used in Wang et al. (2017)) accounting for market risk (beta), stock risk (idiosyncratic return volatility) measured using the Fama-French 3 factor model, and analyst forecast dispersion. We find that our results are robust to the addition of these control variables.[7] Finally, in Table I we repeat our analysis in Table 5 replacing the Fama-Macbeth procedure with the double clustered standard error regression model approach of Thompson (2011). In order to make Table I comparable with our earlier results, we perform a couple of adjustments. Firstly, we decile rank all independent and control variables over the month and use the decile-ranked variables in the regression. This is to ensure that extreme values in one month do not have undue influence on the results (in the Fama-Macbeth procedure, coefficients are averaged across months and so this is less of a concern). Secondly, we perform a weighted regression, where observations within a month are weighted to ensure that each month receives an equal weight (in the Fama-Macbeth procedure each month is equally weighted regardless of the number of observations in the month). We find that the results

---

[5] In addition to the tests mentioned here, we follow recent literature testing the CGO model (An et al., 2019; Wang et al. 2017) and also perform sub-sample tests by excluding NASDAQ stocks or by splitting the sample in half by date and running a regression on the two sub-samples. Our results are robust to these additional tests.

[6] For the January sub-sample, Model A produces a higher average R-squared than models B and C, while for the February-December sub-sample the CGO composite variables once more provide superior explanatory power as is the case in the full-sample analysis. Grinblatt and Han (2005) interpret the negative CGO variable in the January sub-sample in the context of tax loss selling in relation to the purchase price. The composite reference points entering the computation of *CGOCom1* and *CGOCom2* are comprised of additional salient points above and beyond the purchase price, which might account for the relative performance of the composite CGO variables in the January sub-sample.

[7] The inclusion of the analyst forecast dispersion variable leads to a reduction in sample size. To limit sample size reduction, we run the models in Table H with only the first two risk controls (i.e. excluding the analyst forecast dispersion variable), so as to limit sample size reduction. Again, our CGO-composite results (untabulated) remain robust.

using the double cluster regression method are in line with those obtained using the Fama-Macbeth procedure.

3.4 Double Sorted Portfolios- Return Analysis

Grinblatt and Han (2005) also analyze the performance of double-sorted portfolios, sorted by the *CGO* and *Mom* variables, where portfolios are sorted into quintiles by one variable and then by the other. This is to test if a variable has predictive power for returns after being sorted by the other variable. Portfolio sorts are less affected by noise and outliers than regression analysis, due to individual stock diversification across quintile portfolios, and a linear relationship between the sorting variable and dependent variable does not have to be assumed.

As our primary interest is in comparing the predictive power of the traditional CGO variable versus the composite variables, we sort by *CGO* and either *CGOCom1* or *CGOCom2*. Each portfolio is rebalanced every month with stocks within each quintile being equally weighted. The bottom decile of stocks by market cap is excluded from portfolio sorts due to liquidity reasons, as they are for the earlier regression analysis. In Table 6 stocks are first sorted by *CGO* into quintiles and then are further sorted into quintiles by *CGOCom1* and in Table 7 stocks are first sorted by *CGO* and then by *CGOCom2*. The lowest numbered quintile represents the lowest values of the variable in question.

[Insert Table *6*]

[Insert Table 7]

When sorted first by *CGO* and then by *CGOCom1* in Table 6A, the average returns of portfolios increase monotonically with their *CGOCom1* quintile, except for the first quintile of *CGO* (CGO-1). The difference between the first and last quintile (5-1) is always significant, except in the case of the first CGO quintile

20

(CGO-1). Table 6B shows the analysis when stocks are first sorted by *CGOCom1* and then by *CGO* into quintiles. Stocks rarely increase with *CGO* quintile and 4 out of 5 of the difference portfolios have a negative value.

Table 7A and B repeat the analysis for *CGO* and the second composite variable, *CGOCom2*. Except for the first *CGO* quintile (CGO-1), the average returns of portfolios increase monotonically with their *CGOCom2* quintile. The difference between the last and first (5-1) *CGOCom2* quintiles are significant, except for the first *CGO* quintile, ranging from 0.24% to 0.79% per month. Even in the case of the first *CGO* quintile (CGO-1), although there is no significant difference between the average *CGOCom2* quintiles as shown by t-statistic on the spread (5-1) portfolio, the average returns on the first quintile (CGO-1/CGOCom2-1) is comprised of high volatility stocks and thus has a far lower compound return than the fifth quintile (CGO-1/CGOCom2-5). This is reflected in the lower t-statistic of portfolio CGO-1/CGOCom2-1 of 2.15, versus the t-statistic of CGO-1/CGO-Com2-5 of 4.13, even though both portfolios have a similar average return.

In Table 7B stocks are first sorted by *CGOCom2* into quintiles. Within each quintile, stocks are sorted into further quintiles by *CGO*. When first sorting by *CGOCom2* there is not a monotonic relationship between *CGO* and future returns in any of the *CGOCom2* quintiles. The difference between the last and first *CGO* quintiles (5-1) are negative and insignificant, except in the case of the *CGOCom2-5* quintile.

The results of the double sorts suggest that the composite CGO variables are a stronger and more consistent predictor of future returns than *CGO*. The composite variables exhibits a monotonic relationship with returns for all the CGO quintiles, except the first CGO quintile and the composite difference portfolios are positive and significant. The reverse is not true, as CGO quintiles largely exhibit no relationship with future returns, after first sorting by the composite variables. For those interested, a comparison between the two

composite variables: *CGOCom1* and *CGOCom2* is available in the Electronic Companion in Tables J & K. This suggests that neither one of the composite variables dominates the other.

3.5 Regression of one-week Ahead Volume against CGO Variables

Prior literature suggests that trading volume may also be sensitive to the reference point. Kliger and Kudryavtsev (2008) examine the extent to which company-specific events, namely earnings announcements, cause investors to adjust their reference points with a resulting impact on trading behavior. They find that when a stock crosses a new reference point, formed by the price level the day following unanticipated earnings news, it triggers higher trading volume. Huddart et al. (2009) show that weekly volume is strikingly higher when a stock price crosses either the upper or lower limit of its past trading range over the year and suggest that these limits are important reference points for investors. An alternative perspective offered by Baucells et al. (2011, p.512) asserts that increases in volume around historic price highs need not be reflective of such prices impacting investor reference points directly; rather that peak prices tend to be above the reference point for most investors, resulting in increased volume because investors, experiencing paper gains, tend to succumb to the disposition effect and sell.

Following Baucells et al. (2011), if our CGO composites reveal the percentage capital gain or loss from the reference point then we would expect them to be related to future trading volume. Specifically, higher levels of CGO should be associated with higher trading activity as disposition prone investors look to dispose of these stocks, whereas stocks with low CGO should be associated with low volume (Shefrin and Statman, 1985). If our CGO composite variables more accurately represent the reference points of market investors than CGO then we would expect them to have a stronger relationship with trading volume than CGO. By examining the predictive ability of CGO composites comprised of reference points based, amongst other things, on historic highs/lows (*max*/*min*) and 52-week highs/lows (*max52*/*min52*), in the context of volume, we are able to examine the extent to which such prices shift reference points and by doing so impact trading

volume. An increased predictive ability of our CGO composites over the original CGO would provide direct support for the view that salient historic highs/lows shift reference points and impact trading volume.

The following analysis converts our earlier dataset into weekly data (Wednesday to Wednesday) to examine the predictive power of CGO and CGO composites on future volume. The weekly approach mirrors Huddart et al. (2009), although our dataset is far larger. We define the dependent variables as follows: VOL is the average daily number of firm shares traded as a percentage of firm shares outstanding in the observation week and ABNVOL is the residual from firm-by-firm OLS regressions of VOL on market volume, where market volume is measured as the average daily number of shares traded on the exchange where the stock is listed (NASDAQ or NYSE/Amex), expressed as a percentage of the number of shares outstanding for issues listed on that exchange in the observation week. Control variables are also taken from Huddart et al. (2009) and calculated as follows: MAX (MIN) is an indicator variable that takes the value 1 if the closing stock price for the observation week is above (below) the highest (lowest) price attained in the 48-week period ending 20 trading days before the last day of the observation week. DIV and EARNANN are indicator variables taking the value 1 if a dividend record date (from CRSP) or an earnings announcement (from COMPUSTAT), respectively, occurs during the observation week. SDVOL is the annualized standard deviation of stock returns computed from the 126 daily observations prior to the observation week. For $i \in \{1,2,3,4\}$, RET_i is the stock return in week $-i$ relative to the event. RET_5 is the return over weeks -26 to -5, inclusive. The returns are split by sign so the returns regressors are $PRET_i = \max (RET_i, 0)$ and $NRET_i = \min (RET_i, 0)$ for $i \in \{1,2,3,4,5\}$. PRET 2 to 5 and NRET 2 to 5 are omitted in the following regression results for reason of brevity.

[Insert Table 8]

23

In Table 9, we regress raw weekly volume (VOL) against CGO and the CGO composites in Models A, B and C as shown in Equation 9. The results show that while all of the CGO variables are predictive of future trading volume, the CGO composites variables in Models B & C have higher coefficients than CGO and the explanatory power of these models is higher in the form of R-Squared than Model A. This suggests that the composite variables are stronger predictors of raw weekly volume than CGO.

$$VOL_t = \beta_0 + \beta_1 CGO_{t-1} + \beta_2 MAX_t + \beta_3 MIN_t + \beta_4 DIV_t + \beta_5 EARNANN_t + \beta_6 SDVOL_{t-1} + \beta_7 PRET1_t +$$

$$\beta_8 PRET2_t + \beta_9 PRET3_t + \beta_{10} PRET4_t + \beta_{11} PRET5 + \beta_{12} NRET1_t + \beta_{13} NRET2_t + \beta_{14} NRET3_t +$$

$$\beta_{15} NRET4_t + \beta_{16} NRET5 + \varepsilon \tag{9}$$

[Insert Table 9]

This result is mirrored in Models D, E and F that show the relationship between CGO and future abnormal volume (ABNVOL). While all three CGO variables are predictive of future abnormal trading volume, the CGO composites have larger coefficients than CGO and Models E and F have more explanatory power to predict future abnormal volume than Model D.

In summary, the results suggest that future volume is more sensitive to the CGO composite variables than for CGO. This is in line with our earlier analysis on future returns, where we show that 1-month forward returns are more sensitive to the CGO composites than to CGO. Both results suggest that the composite reference points used in CGOCom1 or CGOCom2 are more accurate reference points than the purchase price used in CGO. Given these composite reference points are formed with input from historic highs/lows (*max*/*min*) and 52-week highs/lows (*max52*/*min52*), respectively, we provide direct support to the view that such prices impact investors' reference points, as discussed in Kliger and Kudryavtsev (2008) and Huddart et al. (2009). At the same time, our use of composite CGO measures supports the view in Baucells et al.

(2011) that observed volume increases around historic highs is reflective of the peak price being above the reference point for most investors, thus placing them in a gain situation and so under the influence of the disposition effect.

3.6 Moderation Analysis: Speculative Stocks

In this section, we examine if retail investors are more sensitive to reference point affects than institutional traders. Dhar and Zhu (2006) suggest that non-professionals exhibit a higher disposition effect, which supports the idea that these non-professionals may be more sensitive to reference points than professional institutional traders. If this is the case then we would expect CGO and the CGO composites to have greater predictive power among the stocks that retail investors are more likely to trade in.

Within the Grinblatt and Han (2005) model investors are split into two categories: either rational or PT/MA investors (short for Prospect Theory/Mental Accounting). The PT/MA investors are subject to the disposition effect, and hence drive returns from the abnormal returns from the CGO variable. If retail investors are more likely to be the irrational, PT/MA traders then the predictive power of CGO and the CGO composite variables could be stronger among the more speculative stocks that are more likely to be traded by these retail investors (Han and Kumar, 2013).

We adopt three proxies for speculative characteristics in stocks based on high turnover, small size or high volatility. The categorizing variables are defined as follows: Avgturn (average of daily turnover over the last year), Mrkcap (log of market capitalization), Ivol (daily idiosyncratic volatility over the last year measured using the Fama-French 3 factor model). Table 10 presents Fama-Macbeth regressions that are the same as Equation (8), except that we add three new independent variables: CGO interacted with turnover, CGO interacted with market capitalization, and CGO interacted with idiosyncratic volatility. We repeat for

both CGO composite variables; CGOCom1 and CGOCom2 to produce 9 models in total across the 3 moderating variables.

[Insert Table 10]

The first three models A, B and C report the coefficients on the turnover interaction term. All three of the interaction terms are positive and significant suggesting that high turnover stocks are more likely to be subject to mispricing caused by PT/MA investors. [8] For both market capitalization (models D, E, F) and price volatility (models G, H, I), however, none of the interactions terms are significant. [9] In summary, only the turnover variable acts as a positive moderator of CGO or the CGO composite variables, with no significant effect from the other two proxies.

The pattern of results may have an alternative explanation suggested by the Grinblatt and Han (2005) model itself. The model suggests that a stock's expected return is monotonically increasing in both CGO and high current turnover. This is because high current turnover closes the mispricing caused by PT/MT investors, which shifts the aggregate demand function closer to rational pricing, as earlier investors liquidate their holding. They show that the predictive power of CGO is increased when it is multiplied by the turnover of the current week, in line with the prediction of their model. It is possible therefore that our average turnover variable, that is calculated using the average turnover over the past year, acts as a proxy for the refreshing of the reference point that occurs as older investors are recycled by new ones. Analysis using the current

---

[8] High frequency trading has been shown to impact market performance and price discovery. To further examine the robustness of our results, we also run a sub-sample regression for high frequency trading stocks, which are ranked in the highest 20% quintile of turnover over the prior month. We find that our results in Table 5 are robust within this subset of stocks.
[9] Following Wang et al., (2017), we also ran value-weighted regressions, with weights given by the square root of market capitalisation over the prior month. We find that our results in Table 5 are robust to the change in weighting.

month's turnover as a moderator produces similar results to those presented here using the average daily annual turnover, supporting this idea.

4. Discussion

The literature on reference point adaptation is still developing, but research such as Baucells et al. (2011) has shown that intermediate prices can affect the reference point of the investor. Our first contribution is to expand on these earlier studies, using real stock price graphs over longer horizons, to investigate the likely reference point for an investor in an externally valid setting. Our experimental results suggest the importance of highs and lows, in addition to the purchase and final prices in line with earlier studies such as Baucells et al. (2011), but we also extend previous findings by demonstrating that highs and lows attained during the last 52-weeks play a strong role in reference point formation over longer time periods. This finding is consistent with George and Hwang (2004), who demonstrate that the distance from 52-week high is a key driver of future stock returns. The implication of our results is that further research should be carried out into the 52-week high as a reference point, for example, Bhootra and Hur (2013) have discovered a recency effect that magnifies the impact of a 52-week high. The 52-week low has received scant attention in the literature and is also worthy of further investigation. Arkes et al. (2008) and Chen and Rao (2002) suggest that reference point adaptation is greater in gains than losses. This asymmetric adjustment may be caused by the greater saliency of recent highs than recent lows, as we demonstrate in our experiment.

Our second contribution is to take the insights we learn from the controlled conditions of the experiment and apply these to the rich and complex setting of the US stock market, using the CGO model developed by Grinblatt and Han (2005). We show that the purchase price is not the only reference point that is predictive of one month ahead returns when plugged into the CGO model. In fact, alternative CGO variables based on the maximum, minimum, 52-week maximum or 52-week minimum are equally good predictors.

This may seem a surprising result given the common use of a purchase price based reference point in financial models such as the CGO, but demonstrates the benefits of exploring more complex models in a controlled environment and then applying in the real-world context.

Thirdly, we create two CGO-Composite variables formed from weighting different salient points in the stock price path, using coefficients determined in the experiment. CGOCom1 is created using the purchase, maximum and minimum prices, while CGOCom2 includes the purchase, 52-week maximum and minimum prices. The traditional CGO variable is no longer a positive predictor of returns when either of the CGO-Composite variables are included in the regression. We also conduct double sorts by both CGO and CGO composites and determine that CGO is rarely predictive of returns after stocks are first sorted by the composites, but the composites are generally predictive of returns even if stocks are first sorted by CGO.

Fourthly, we find that the CGO Composite variables have a stronger relationship with future trading volume than CGO. Both 1-week ahead raw volume and abnormal volume are more sensitive to changes in the CGO Composites than to changes in CGO. The results suggest that our CGO Composite variables, formed from a composite of salient points predicted in the experiment, are better predictors of forward returns and forward volume than the traditional CGO variable. The implication is that reference points are formed from multiple salient points in the stock price path, rather than the purchase price alone, and thus models using these points have additional explanatory power. Finally, our results are robust to investor segment analysis, using proxies for speculative stocks to measure the influence of CGO across different groups of investors.

Our results have implications both for academics interested in reference point adjustment and investment professionals who wish to study how reference points cause mispricing in markets. The distortions in market prices caused by reference points, which we demonstrate using the market data model, lead to profitable arbitrage opportunities which could be capitalised upon by Investment Managers.

Looking to future research, many other financial models assume a fixed, purchase price based reference point such as realization utility (Barberis and Xiong, 2012), and our results suggest that the impact of adjusting reference points in other financial models is worthy of investigation. The disposition effect is a popular area of study, for example, and yet the original assumption of Shefrin and Statman (1985) regarding the reference point has largely remained in place. Reference point adjustment is a subjective process that is challenging to model, being subject to many variables. The response to these variables may vary by individual characteristics (Dhar and Zhu, 2006), environmental factors, and emotions (Summers and Duxbury, 2012). Investigation of the influence of these additional factors could provide new insights into investor behavior and allow the potential development of more accurate reference point predictions.

Acknowledgements

References

AN, L., WANG, H., WANG, J. & YU, J. 2019. Lottery-related anomalies: the role of reference-dependent preferences. *Management Science, Forthcoming*.

ANDERSON, A. M. & DYL, E. A. 2007. Trading Volume: NASDAQ and the NYSE. *Financial Analysts Journal,* 63**,** 79-86.

ARKES, H. R., HIRSHLEIFER, D., JIANG, D. & LIM, S. 2008. Reference point adaptation: Tests in the domain of security trading. *Organizational Behavior and Human Decision Processes,* 105**,** 67-81.

BAMBERGER, P. & FIEGENBAUM, A. 1996. The role of strategic reference points in explaining the nature and consequences of human resource strategy. *Academy of Management Review,* 21**,** 926-958.

BARBERIS, N., SHLEIFER, A. & VISHNY, R. 1998. A model of investor sentiment. *Journal of Financial Economics,* 49**,** 307-343.

BARBERIS, N. & XIONG, W. 2012. Realization utility. *Journal of Financial Economics,* 104**,** 251-271.

BAUCELLS, M., WEBER, M. & WELFENS, F. 2011. Reference-point formation and updating. *Management Science,* 57**,** 506-519.

BEN-DAVID, I. & HIRSHLEIFER, D. 2012. Are investors really reluctant to realize their losses? Trading responses to past returns and the disposition effect. *Review of Financial Studies,* 25**,** 2485-2532.

BHATTACHARYA, U., HOLDEN, C. W. & JACOBSEN, S. 2012. Penny wise, dollar foolish: Buy–sell imbalances on and around round numbers. *Management Science,* 58**,** 413-431.

BHOOTRA, A. & HUR, J. 2013. The timing of 52-week high price and momentum. *Journal of Banking & Finance,* 37**,** 3773-3782.

CACIOPPO, J. T. & BERNTSON, G. G. 1994. Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin,* 115**,** 401.

CHEN, H. & RAO, A. R. 2002. Close encounters of two kinds: False alarms and dashed hopes. *Marketing Science,* 21**,** 178-196.

DA, Z., LIU, Q. & SCHAUMBURG, E. 2013. A closer look at the short-term return reversal. *Management Science,* 60**,** 658-674.

DHAR, R. & ZHU, N. 2006. Up close and personal: Investor sophistication and the disposition effect. *Management Science,* 52**,** 726-740.

FAMA, E. F. & MACBETH, J. D. 1973. Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy***,** 607-636.

GEORGE, T. J. & HWANG, C.-Y. 2004. The 52-week high and momentum investing. *The Journal of Finance,* 59**,** 2145-2176.

GNEEZY, U. 2005. Updating the reference level: Experimental evidence. *Experimental business research.* Springer.

GRINBLATT, M. & HAN, B. 2005. Prospect theory, mental accounting, and momentum. *Journal of Financial Economics,* 78**,** 311-339.

GROSSHANS, D. & ZEISBERGER, S. 2017. All's Well that Ends Well? On the Importance of How Returns are Achieved. *Journal of Banking & Finance,* Forthcoming.

HAN, B. & KUMAR, A. 2013. Speculative retail trading and asset prices. *Journal of Financial and Quantitative Analysis,* 48**,** 377-404.

HEATH, C., HUDDART, S. & LANG, M. 1999. Psychological factors and stock option exercise. *Quarterly Journal of Economics***,** 601-627.

HEYMAN, J., MELLERS, B., TISHCENKO, S. & SCHWARTZ, A. 2004. I was pleased a moment ago: How pleasure varies with background and foreground reference points. *Motivation and Emotion,* 28**,** 65-83.

HOFFMANN, A. O., HENRY, S. F. & KALOGERAS, N. 2013. Aspirations as reference points: an experimental investigation of risk behavior over time. *Theory and Decision***,** 1-18.

HUDDART, S., LANG, M. & YETMAN, M. H. 2009. Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence. *Management Science,* 55**,** 16-31.

JEGADEESH, N. & TITMAN, S. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance,* 48**,** 65-91.

KAHNEMAN, D. & TVERSKY, A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society***,** 263-291.

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

KAUSTIA, M. 2010. Prospect Theory and the Disposition Effect. *Journal of Financial and Quantitative Analysis,* 45**,** 791-812.

KLIGER, D. & KUDRYAVTSEV, A. 2008. Reference point formation by market investors. *Journal of Banking & Finance,* 32**,** 1782-1794.

NEWEY, W. K. & WEST, K. D. 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica,* 55**,** 703-708.

POTESHMAN, A. M. & SERBIN, V. 2003. Clearly Irrational Financial Market Behavior: Evidence from the Early Exercise of Exchange Traded Stock Options. *The Journal of Finance,* 58**,** 37-70.

RABIN, M. & VAYANOS, D. 2010. The gambler's and hot-hand fallacies: Theory and applications. *The Review of Economic Studies,* 77**,** 730-778.

ROSENBERG, B. & RUDD, A. 1982. Factor Related and Specific Returns of Common Stocks: Serial Correlation and Market Inefficiency. *The Journal of Finance,* 37**,** 543-554.

SAMUELSON, W. & ZECKHAUSER, R. 1988. Status quo bias in decision making. *Journal of Risk and Uncertainty,* 1**,** 7-59.

SHEFRIN, H. & STATMAN, M. 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance***,** 777-790.

SUMMERS, B. & DUXBURY, D. 2012. Decision-dependent emotions and behavioral anomalies. *Organizational Behavior and Human Decision Processes,* 118**,** 226-238.

THOMPSON, S. B. 2011. Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics,* 99**,** 1-10.

TIBSHIRANI, R. & HASTIE, T. 2007. Outlier sums for differential gene expression analysis. *Biostatistics,* 8**,** 2-8.

WANG, H., YAN, J. & YU, J. 2017. Reference-dependent preferences and the risk–return trade-off. *Journal of Financial Economics,* 123**,** 395-414.

WHYTE, G. 1986. Escalating commitment to a course of action: A reinterpretation. *Academy of Management Review,* 11**,** 311-321.

Appendix

Instructions

Survey Instructions

Assume that you purchased a stock in a particular month, which we will label as month 0 in the graphs that follow. Since then you have monitored the share price closely. Today, you are considering selling the stock.

Your task will be to indicate the selling price at which you would feel neutral (i.e. feel neither predominantly positive nor negative) about selling the stock.

On the following screens, you will be shown a series of 30 share price graphs. The graphs will range in duration from 6 months to 5 years. Assume that you purchased the stock in a particular month, which we will label month 0 in the graphs that follow, and you are now considering selling the stock at the latest point shown on the graph.

This task is not about your maths skills and there are no right or wrong answers. We are just interested in your personal opinion, so just answer as honestly as you can giving an intuitive price.

A progress bar will show your progress through the survey.

You will be able re-read these instructions in the experiment.

Example Chart

At what sale price would you feel neither predominantly positive nor negative about selling?

Choose the price for which you would feel exactly neutral by entering it in the text box below.

Click to re-read Survey Instructions

Statistical Tables

Table 1: Summary of Reference Points and Price Sequences

Chart Number	Mean Reference	Median Reference	SD Reference	Purchase	Maximum	Minimum	Average	Final
1	13.14	13	0.96	13.33	14.58	10.67	13.08	11.83
2	3.09	3	0.34	3.38	3.5	2.38	2.9	2.38
3	8.87	9	1.33	7	10.3	7	8.64	10.29
4	13.51	14	1.73	13.5	15.25	8.25	11.04	15.25
5	6.11	6	0.61	5.9	6.72	4.76	5.69	6.6
6	123.9	125	42.55	66.13	198.19	66.13	128.26	155
7	21.64	22	1.11	22	23.45	19.7	21	20.75
8	8.23	8	1.17	9.41	9.77	5.18	6.9	6.85
9	22.29	22	2.85	18.53	25.97	18.29	22.42	25.31
10	7.12	7	1.23	6.05	8.9	3.05	5.96	8.15
11	12.95	13	1.89	15	15.5	9.5	11.29	10
12	8.68	9	1.55	5.88	10.75	5.5	8.81	10
13	5.27	5	1.38	5.63	8	2.5	4.85	3.06
14	11.39	12	2.58	14.5	14.5	6.5	9.22	7
15	18.55	20	5.15	20.13	29.38	9.25	17.64	10.5
16	16.57	16	4.37	9.5	22.06	9.25	14.34	21.71
17	6.52	6.5	0.82	6.88	8	5	6.46	5.5
18	32.93	32	4.06	31.38	41.25	21.75	30.4	31.13
19	40.22	40	21.57	10	73.06	7.38	21.64	62.54
20	23.13	22	5.87	17.7	35.17	17.2	26	18.49
21	90.74	95	24.35	54.04	126.76	52	78.41	121.36
22	14.85	15	1.75	14.25	18.88	10.38	13.99	15
23	21.46	20	4.75	26.92	28.22	10.55	18.79	14.88
24	78.98	80	31.78	78.38	151	22	84.08	24.5
25	2.97	3	0.8	1.91	4.66	1.25	2.61	2.78
26	27.08	30	6.63	30.94	34.19	2.36	13.05	26.25
27	48.7	50	5.58	46.5	58.13	25.88	40.87	49.88
28	35.13	34.5	7.54	34.22	50.16	14.1	28.17	37.26
29	33.04	35	11.94	13.49	56.03	12.9	31.91	38.16
30	17.74	18	3.09	15.51	22.74	8.03	14.34	19.81

This table reports summary statistics for the charts we use in the experiment. Chart number 1-6=6 months, 7-12=12 months, 13-18= 2 years, 19-24 = 3 years, 25-30= 5 years. Reference is the mean or median reference point provided by participants across the chart. Purchase is the purchase price of the chart. Maximum is the maximum price of the chart. Minimum is the minimum price of the chart. Average is the average price of a chart. Final is the final price of a chart.

Table 2: Regression Analysis using Price Variables

VARIABLES	Model A	Model B	Model C	Model D	Model E	Model F
Purchase	0.328***	0.328***	0.341***	0.435***	0.434***	0.431***
	(10.03)	(9.660)	(9.812)	(14.09)	(16.74)	(11.53)
Maximum	0.294***	0.292***	0.235***			
	(6.430)	(11.85)	(10.05)			
Minimum	0.138	0.135***				
	(1.847)	(4.326)				
Average	-0.00280		0.114***	-0.00286		
	(-0.0300)		(2.914)	(-0.0582)		
Final	0.232***	0.232***	0.265***	0.00440	0.00554	
	(6.768)	(7.373)	(9.420)	(0.0914)	(0.126)	
Max52				0.446***	0.444***	0.448***
				(9.903)	(12.40)	(17.59)
Min52				0.109***	0.109***	0.112***
				(3.300)	(4.077)	(4.042)
Constant	0.168	0.171	0.314***	0.101	0.103	0.116
	(1.430)	(1.459)	(3.212)	(0.994)	(0.948)	(0.702)
Observations	5,005	5,005	5,005	5,005	5,005	5,005
R^2	0.837	0.837	0.837	0.836	0.836	0.836
Adj R^2	0.837	0.837	0.837	0.836	0.836	0.836

This table reports results for predictive regressions of reference points on a set of salient prices. Dependent variable= reference point provided by participant. Purchase= purchase price shown in chart. Maximum= maximum price shown in chart. Minimum= minimum price shown in chart. Average= arithmetic average of prices shown in chart. Final= final price shown in chart. Max52=52-week high price shown in chart for charts of 12 months or longer; 6 month high otherwise. Min52=52-week low price shown in chart for charts of 12 months or longer; 6 month low otherwise. Robust t-statistics in parentheses, clustered by participant- *** p<0.01, ** p<0.05.

Table 3: VIF Analysis

Variable	Model A	Model B	Model C	Model D	Model E	Model F
Purchase	7.46	7.34	6.68	6.57	4.96	2.92
Maximum	126.20	12.40	25.70			
Minimum	53.87	9.46				
Average	177.50		31.18	50.38		
Final	23.77	11.39	5.43	40.47	28.56	
Max52				90.68	25.99	6.17
Min52				9.63	8.43	4.23
Mean VIF	77.76	10.15	17.25	39.62	16.98	4.44

Notes: VIF analysis shown for models in Table 2.

VARIABLES	Model A	Model B	Model C	Model D	Model E	Model F
CGO	0.00511***					
	(4.429)					
CGOMax		0.00384***				
		(4.111)				
CGOMin			0.0202***			
			(5.208)			
CGOAverage				0.0167***		
				(6.305)		
CGOMax52					0.0108***	
					(5.710)	
CGOMin52						0.0189***
						(3.615)
Mom	0.00491***	0.00565***	0.00481***	0.00292	0.00485***	0.00502***
	(2.934)	(3.500)	(2.611)	(1.797)	(3.223)	(2.654)
STR	-0.0541***	-0.0542***	-0.0551***	-0.0587***	-0.0570***	-0.0554***
	(-10.01)	(-9.860)	(-9.977)	(-10.92)	(-10.25)	(-9.685)
LTR	-0.000959	-0.000769	-0.000876	-0.00106	-0.000755	-0.000518
	(-1.736)	(-1.483)	(-1.447)	(-1.798)	(-1.288)	(-0.915)
Avgturn	-0.673**	-0.777**	-0.579	-0.559	-0.650	-0.752**
	(-2.076)	(-2.285)	(-1.572)	(-1.503)	(-1.728)	(-2.238)
Mrkcap	-0.000279	-0.000344	0.000172	-0.000316	-0.000518	0.000126
	(-0.649)	(-0.818)	(0.395)	(-0.727)	(-1.232)	(0.297)
BM	0.00205***	0.00185**	0.00253***	0.00225***	0.00174**	0.00237***
	(2.682)	(2.445)	(3.685)	(3.027)	(2.325)	(3.431)
Constant	0.0144***	0.0159***	0.00697**	0.0145***	0.0175***	0.00887***
	(4.166)	(4.768)	(2.213)	(4.166)	(5.165)	(2.834)
Observations	1,749,966	1,749,966	1,749,966	1,749,966	1,749,966	1,749,966
Number of groups	647	647	647	647	647	647
Average R^2	0.0684	0.0685	0.0656	0.0677	0.0690	0.0663

Notes: This table reports results for predictive Fama and MacBeth (1973) regressions of one-month ahead returns on CGO and a set of control variables. Dependent variable= 1 month return in month t+1. CGO calculated using turnover adjusted purchase price as shown in Equation 5. CGOMax calculated as per CGO but replacing the purchase price with the maximum price. CGOMin calculated as per CGO but replacing the purchase price with the minimum price. CGOAverage calculated as per CGO but replacing the purchase price with the average price. CGOMax52 calculated as per CGO but replacing the purchase price with the 52-week maximum price. CGOMin52 calculated as per CGO but replacing the purchase price with the 52-week minimum price. MOM=12month momentum, excluding the last month. STR=Short term reversal, calculated as the return of the last month t. LTR=Long-term reversal, calculated as the return over the last 3 years excluding the last year. AvgTurn is average daily turnover over last year. Mrkcap is the log of market capitalisation in thousands. BM is log of the book to market ratio. T-statistics in parentheses are Newey-West adjusted.
*** p<0.01, ** p<0.05.

Table 5: Regression using CGO Composite Variables

VARIABLES	Model A	Model B	Model C	Model D
CGO	0.00511***	-0.0327***	-0.0103***	
	(4.429)	(-6.450)	(-3.633)	
CGOCom1		0.0562***		0.00326
		(7.414)		(0.566)
CGOCom2			0.0261***	0.00929
			(5.663)	(1.552)
Mom	0.00491***	0.00553***	0.00384**	0.00398***
	(2.934)	(3.351)	(2.428)	(2.586)
STR	-0.0541***	-0.0526***	-0.0587***	-0.0569***
	(-10.01)	(-9.665)	(-10.35)	(-10.27)
LTR	-0.000959	-0.000134	-0.000547	-0.000973
	(-1.736)	(-0.230)	(-0.923)	(-1.766)
Avgturn	-0.673**	-0.857***	-0.498	-0.533
	(-2.076)	(-2.601)	(-1.507)	(-1.509)
Mrkcap	-0.000279	-0.000514	-0.000490	-0.000450
	(-0.649)	(-1.190)	(-1.160)	(-1.072)
BM	0.00205***	0.00139	0.00178**	0.00198***
	(2.682)	(1.848)	(2.398)	(2.634)
Constant	0.0144***	0.0178***	0.0166***	0.0158***
	(4.166)	(5.205)	(4.864)	(4.668)
Observations	1,749,966	1,749,966	1,749,966	1,749,966
Average R^2	0.0683	0.0710	0.0722	0.0721
Adj R^2	0.0638	0.0659	0.0671	0.0670
Number of groups	647	647	647	647

Notes: This table reports results for predictive Fama and MacBeth (1973) regressions of one-month ahead returns on CGO with CGO Composites and a set of control variables. Dependent variable= 1 month return in month t+1. CGO calculated using turnover adjusted purchase price as shown in Equation 5. CGOCom1 calculated as per CGO but replacing the purchase price with Refcom1, calculated as per Equation 6. CGOCom2 calculated as per CGO but replacing the purchase price with Refcom2, calculated as per Equation 7. MOM=12month momentum, excluding the last month. STR=Short term reversal, calculated as the return of the last month t. LTR=Long-term reversal, calculated as the return over the last 3 years excluding the last year. AvgTurn is average daily turnover over last year. Mrkcap is the log of market capitalisation in thousands. BM is log of the book to market ratio. T-statistics in parentheses are Newey-West adjusted- *** p<0.01, ** p<0.05.

Table 6A: Double Portfolio Sorts by CGO and CGO-Com1

	CGO-1	CGO-2	CGO-3	CGO-4	CGO-5
CGOCom1-1	0.76	0.59	0.78	0.90	1.07
	1.76	2.21**	3.28**	3.87**	4.97**
CGOCom1-2	0.70	0.87	1.14	1.23	1.35
	1.93	3.46**	5.15**	5.90**	6.64**
CGOCom1-3	0.82	1.00	1.19	1.31	1.63
	2.52**	4.15**	5.53**	6.69**	7.69**
CGOCom1-4	0.94	1.05	1.29	1.33	1.73
	3.13**	4.34**	6.21**	6.64**	7.96**
CGOCom1-5	1.09	1.27	1.42	1.52	2.09
	3.89**	5.53**	7.34**	7.42**	8.28**
5-1	**0.33**	**0.68**	**0.64**	**0.62**	**1.03**
	1.52	**6.06****	**5.62****	**5.48****	**6.87****

Notes: This table (Panels A & B) reports returns in double sorted portfolios based on values of CGO and CGOCom1. At the end of each month, stocks are sorted into 5 portfolios by CGO and CGOCom1. Stocks in a portfolio are equally weighted. Each portfolio is held for one month and the time series average return in reported in monthly percent. Newey-West corrected t-Statistics are shown below performance.

**=significance at the 5% level.

Table 6B: Double Portfolio Sorts by CGO-Com1 and CGO

	CGOCom1-1	CGOCom1-2	CGOCom1-3	CGOCom1-4	CGOCom1-5
CGO-1	0.86	1.05	1.30	1.41	1.55
	2.00**	4.14**	6.08**	7.36**	7.47**
CGO-2	0.82	0.98	1.27	1.27	1.52
	2.25**	4.00**	5.85**	6.70**	7.53**
CGO-3	0.79	0.98	1.24	1.30	1.54
	2.42**	3.91**	5.87**	6.18**	7.35**
CGO-4	0.94	0.91	1.11	1.29	1.69
	3.20**	3.78**	5.27**	6.33**	7.31**
CGO-5	0.58	0.71	0.88	1.13	1.92
	2.03**	2.83**	3.67**	5.18**	7.90**
5-1	**-0.28**	**-0.34**	**-0.42**	**-0.27**	**0.37**
	-1.29	**-2.83****	**-3.66****	**-2.70****	**3.16****

Table 7A: Double Sorts by CGO & CGOCom2

	CGO-1	CGO-2	CGO-3	CGO-4	CGO-5
CGOCom2-1	0.97	0.65	0.97	1.05	1.27
	2.15**	2.23**	3.70**	4.06**	5.28**
CGOCom2-2	0.83	0.86	1.12	1.32	1.45
	2.13**	3.16**	4.47**	5.75**	6.68**
CGOCom2-3	0.83	1.05	1.22	1.34	1.61
	2.49**	3.97**	5.23**	6.44**	7.40**
CGOCom2-4	0.96	1.06	1.31	1.39	1.79
	3.09**	4.25**	5.92**	6.44**	7.62**
CGOCom2-5	1.21	1.33	1.37	1.47	2.05
	4.13**	5.62**	6.63**	7.08**	7.74**
5-1	**0.24**	**0.68**	**0.40**	**0.43**	**0.79**
	0.97	**4.42****	**2.74****	**3.03****	**4.89****

Notes: This table (Panels A & B) reports returns of double sorted portfolios based on values of CGO and CGOCom2. At the end of each month, stocks are sorted into 5 portfolios by CGO and CGOCom2. Stocks in a portfolio are equally weighted. Each portfolio is held for one month and the time series average return in reported in monthly percent. Newey-West corrected t-Statistics are shown below performance.
**=significance at the 5% level.

Table 7B: Double Sorts by CGOCom2 & CGO

	CGOCom2-1	CGOCom2-2	CGOCom2-3	CGOCom2-4	CGOCom2-5
CGO-1	1.08	1.17	1.33	1.37	1.59
	2.41**	4.35**	5.67**	6.64**	7.27**
CGO-2	0.93	1.07	1.28	1.26	1.54
	2.48**	4.07**	5.52**	6.45**	7.29**
CGO-3	0.81	0.81	1.18	1.34	1.57
	2.41**	3.01**	5.08**	6.27**	7.06**
CGO-4	0.93	0.95	1.15	1.39	1.78
	2.94**	3.49**	4.83**	6.27**	7.38**
CGO-5	0.66	0.93	1.10	1.30	1.97
	2.14**	3.47**	4.26**	5.37**	7.73**
5-1	**-0.42**	**-0.25**	**-0.23**	**-0.07**	**0.38**
	-1.73	**-1.58**	**-1.47**	**-0.55**	**3.06****

Table 8: Descriptive Statistics for Volume Analysis

	VOL	ABNVOL	MAX	MIN	DIV	EARNANN	SDVOL	PRET1	NRET1
Mean	0.317	0.003	0.159	0.110	0.035	0.048	0.484	0.026	-0.021
SD	0.595	0.003	0.337	0.272	0.178	0.172	0.282	0.119	0.035
Median	0.180	0.002	0.001	0.020	0.000	0.000	0.424	0.005	-0.006
Max	18.036	0.043	1.000	0.999	1.000	0.846	4.649	5.549	0.000
Min	0.000	-0.007	0.000	0.000	0.000	0.000	0.036	0.000	-0.453
Skew	11.516	2.609	2.607	3.838	5.668	6.152	3.218	12.557	-3.730

Notes: This table reports summary statistics for variables used in volume analysis. All number presented are the time-series average of the cross-sectional statistics. VOL is the average daily number of firm shares traded as a percentage of firm shares outstanding in the observation week. ABNVOL is the residual from firm-by-firm OLS regressions of VOL on market volume, where market volume is measured as the average daily number of shares traded on the exchange where the stock is listed (NASDAQ or NYSE/Amex), expressed as a percentage of the number of shares outstanding for issues listed on that exchange in the observation week. MAX (MIN) is an indicator variable that takes the value 1 if the closing stock price for the observation week is above (below) the highest (lowest) price attained in the 48-week period ending 20 trading days before the last day of the observation week. DIV and EARNANN are indicator variables taking the value 1 if a dividend record date (from CRSP) or an earnings announcement (from COMPUSTAT), respectively, occurs during the observation week. SDVOL is the annualized standard deviation of stock returns computed from the 126 daily observations prior to the observation week. For $i \in \{1,2,3,4\}$, RETi is the stock return in week $-i$ relative to the event. RET5 is the return over weeks -26 to -5, inclusive. The returns are split by sign so the returns regressors are $PRETi=\max(RETi,0)$ and $NRETi=\min(RETi,0)$ for $i \in \{1,2,3,4,5\}$.

Table 9: Weekly Regression of Volume on CGO and Control Variables

VARIABLES	Model A	Model B	Model C	Model D	Model E	Model F
CGO	0.0743***			0.000297***		
	(7.022)			(4.675)		
CGOCom1		0.153***			0.000733***	
		(8.151)			(6.614)	
CGOCom2			0.171***			0.000829***
			(8.104)			(6.606)
MAX	0.141***	0.140***	0.139***	0.000163***	0.000155***	0.000156***
	(11.86)	(11.78)	(11.72)	(4.778)	(4.582)	(4.500)
MIN	0.0603***	0.0640***	0.0688***	- 0.000189***	-0.000165***	- 0.000139***
	(9.632)	(10.05)	(10.17)	(-7.106)	(-6.187)	(-5.255)
DIV	0.00915***	0.00850***	0.00848***	- 0.000164***	-0.000168***	- 0.000169***
	(3.456)	(3.286)	(3.309)	(-8.301)	(-8.617)	(-8.705)
EARNANN	0.132***	0.132***	0.132***	7.50e-05***	7.76e-05***	7.76e-05***
	(13.17)	(13.17)	(13.16)	(6.436)	(6.646)	(6.661)
SDVOL	0.121***	0.135***	0.140***	0.000836***	0.000926***	0.000950***
	(8.753)	(9.507)	(9.653)	(9.965)	(11.02)	(11.22)
PRET1	1.593***	1.646***	1.630***	0.00131***	0.00159***	0.00151***
	(15.60)	(15.68)	(15.64)	(13.09)	(13.23)	(13.59)
NRET1	-1.290***	-1.245***	-1.277***	-0.00252***	-0.00231***	-0.00246***
	(-15.70)	(-15.85)	(-15.84)	(-14.46)	(-14.32)	(-14.59)
Constant	0.0754***	0.0721***	0.0679***	0.00211***	0.00208***	0.00205***
	(10.35)	(9.910)	(9.734)	(19.61)	(19.52)	(19.83)
Observations	10,123,143	10,123,143	10,123,143	10,134,527	10,134,527	10,134,527
Number of groups	2,687	2,687	2,687	2,687	2,687	2,687
Average R^2	0.139	0.141	0.142	0.108	0.109	0.110

Notes: This table reports results for predictive Fama and MacBeth (1973) regressions of VOL (Models A, B, C) or ABNVOL (Models D, E, F) on CGO, CGO Composites, and a set of control variables. Refer to Table 8 for definitions of dependent and control variables. CGO calculated using turnover adjusted purchase price as shown in Equation 5. CGOCom1 calculated as per CGO but replacing the purchase price with Refcom1, calculated as per Equation 6. CGOCom2 calculated as per CGO but replacing the purchase price with Refcom2, calculated as per Equation 7. All CGO variables are lagged by one week relative to the dependent variables. T-statistics in parentheses are Newey-West adjusted- *** p<0.01, ** p<0.05.

VARIABLES	Model A	Model B	Model C	Model D	Model E	Model F	Model G	Model H	Model I
Avgturn*CGO	4.511***								
	(2.858)								
Avgturn*CGOCom1		7.390***							
		(2.911)							
Avgturn**CGOCom222			6.445***						
			(2.661)						
Mrkcap*CGO				0.000611					
				(1.107)					
Mrkcap*CGOCom1					0.00103				
					(1.181)				
Mrkcap*CGOCom2						0.000822			
						(1.013)			
Ivol*CGO							0.106		
							(1.018)		
Ivol*CGOCom1								0.232	
								(1.412)	
Ivol*CGOCom2									0.267
									(1.718)
CGO	0.00108			0.00489**			0.00243		
	(0.727)			(1.975)			(1.005)		
CGOCom1		0.00214			0.00894**			0.00437	
		(0.887)			(2.291)			(1.151)	
CGOCom2			0.00261			0.00943***			0.00412
			(0.926)			(2.684)			(1.046)
Mom	0.00405**	0.00384**	0.00375**	0.00442***	0.00397***	0.00378***	0.00571***	0.00519***	0.00486***
	(2.488)	(2.436)	(2.476)	(2.822)	(2.584)	(2.637)	(3.801)	(3.497)	(3.417)
STR	-0.0557***	-0.0557***	-0.0573***	-0.0551***	-0.0556***	-0.0575***	-0.0547***	-0.0552***	-0.0568***
	(-10.15)	(-10.05)	(-10.39)	(-10.28)	(-10.30)	(-10.66)	(-10.19)	(-10.18)	(-10.43)
LTR	-0.00110**	-0.00110**	-0.00105**	-0.00103	-0.00107**	-0.00102	-0.000986	-0.00106**	-0.00109**
	(-2.167)	(-2.195)	(-1.993)	(-1.899)	(-2.047)	(-1.844)	(-1.789)	(-2.062)	(-2.033)
Avgturn	-0.856***	-0.665**	-0.592	-0.522	-0.567	-0.530	-0.431	-0.494	-0.501
	(-2.869)	(-2.181)	(-1.694)	(-1.770)	(-1.736)	(-1.516)	(-1.797)	(-1.834)	(-1.680)
Mrkcap	-0.000382	-0.000449	-0.000474	-0.000366	-0.000413	-0.000464	-0.000706**	-0.000711**	-0.000708**
	(-0.892)	(-1.053)	(-1.119)	(-0.858)	(-0.961)	(-1.071)	(-2.235)	(-2.259)	(-2.250)
BM	0.00208***	0.00203***	0.00196***	0.00213***	0.00208***	0.00203***	0.00166**	0.00171***	0.00169**
	(2.730)	(2.677)	(2.586)	(2.824)	(2.780)	(2.700)	(2.552)	(2.619)	(2.580)
Ivol							-0.138**	-0.109	-0.0881
							(-2.057)	(-1.603)	(-1.232)
Constant	0.0154***	0.0160***	0.0162***	0.0146***	0.0154***	0.0158***	0.0188***	0.0188***	0.0185***
	(4.457)	(4.670)	(4.764)	(4.204)	(4.474)	(4.576)	(6.742)	(6.732)	(6.625)
Observations	1,749,966	1,749,966	1,749,966	1,749,966	1,749,966	1,749,966	1,749,177	1,749,177	1,749,177
Number of groups	647	647	647	647	647	647	647	647	647
Average R^2	0.0725	0.0727	0.0730	0.0721	0.0723	0.0732	0.0785	0.0784	0.0791

Notes: This table reports results for predictive Fama and MacBeth (1973) regressions of one-month ahead returns on CGO and a set of control variables. Dependent variable= 1 month return in month t+1. CGO calculated using turnover adjusted purchase price as shown in Equation 5. CGOCom1 calculated as per CGO but replacing the purchase price with Refcom1, calculated as per Equation 6. CGOCom2 calculated as per CGO but replacing the purchase price with Refcom2, calculated as per Equation 7. MOM=12month momentum, excluding the last month. STR=Short term reversal, calculated as the return of the last month t. LTR=Long-term reversal, calculated as the return over the last 3 years excluding the last year. AvgTurn is average daily turnover over last year. Mrkcap is the log of market capitalisation in thousands. BM is log of the book to market ratio. Ivol is daily idiosyncratic volatility calculated over 12 months using the 3 factor Fama-French model. T-statistics in parentheses are Newey-West adjusted- *** p<0.01, ** p<0.05.

Electronic Companion- Additional Statistical Tables

Table A- Regression of Reference Point against Independent Variables

VARIABLES	Model A	Model B	Model C	Model D	Model E	Model F
Purchase	0.344***	0.343***	0.350***	0.447***	0.440***	0.416***
	(11.06)	(10.49)	(10.32)	(14.81)	(17.91)	(11.40)
Maximum	0.254***	0.271***	0.225***			
	(4.801)	(12.00)	(9.429)			
Minimum	0.0674	0.103***				
	(0.767)	(3.110)				
Average	0.0364		0.0929**	-0.0222		
	(0.333)		(2.148)	(-0.397)		
Final	0.264***	0.256***	0.280***	0.0411	0.0498	
	(7.816)	(8.590)	(10.52)	(0.818)	(1.172)	
Max52				0.427***	0.411***	0.445***
				(8.679)	(12.48)	(18.20)
Min52				0.0842**	0.0792**	0.114***
				(2.134)	(2.516)	(4.036)
Constant	0.367**	0.322**	0.437***	0.252	0.271	0.393
	(2.023)	(1.995)	(3.030)	(1.631)	(1.709)	(1.855)
Observations	5,033	5,033	5,033	5,033	5,033	5,033
R^2	0.826	0.826	0.826	0.826	0.826	0.826
Adj R^2	0.826	0.826	0.826	0.826	0.826	0.826

This table reports results for predictive regressions of reference points on a set of salient prices. Data is cleaned using the outlier sum method of Tibshirani and Hastie, 2007. The median reference point and MAD are calculated separately for each of the 30 charts. The Outlier Sum screening method removes a total of 697 reference points from the raw data, which contains 5730 reference points before exclusions. Dependent variable= reference point provided by participant. Purchase= purchase price shown in chart. Maximum= maximum price shown in chart. Minimum= minimum price shown in chart. Average= arithmetic average of prices shown in chart. Final= final price shown in chart. Max52=52-week high price shown in chart for charts of 12 months or longer; 6-month high otherwise. Min52=52-week low price shown in chart for charts of 12 months or longer; 6-month low otherwise. Robust t-statistics in parentheses, clustered by participant *** p<0.01, ** p<0.05.

Table B- VIF Analysis

Variable	Model A	Model B	Model C	Model D	Model E	Model F
Purchase	7.80	7.67	6.94	7.25	5.15	2.92
Maximum	129.58	12.79	25.82			
Minimum	56.50	9.83				
Average	181.86		31.65	51.01		
Final	25.18	11.79	5.47	41.47	29.59	
Max52				91.61	26.93	6.14
Min52				9.76	8.55	4.18
Mean VIF	80.18	10.52	17.47	40.22	17.55	4.41

Notes: VIF analysis for models shown in Table A.

Table C- Regression of CGO against Independent CGO Variables

VARIABLES	Model A	Model B	Model C	Model D	Model E	Model F
CGOPurchase	0.300***	0.291***	0.311***	0.362***	0.469***	0.353***
	(10.12)	(9.687)	(10.88)	(13.60)	(20.09)	(13.10)
CGOMax	0.251***	0.309***	0.210***			
	(8.421)	(12.35)	(10.01)			
CGOMin	0.0686	0.151***				
	(1.843)	(7.689)				
CGOAverage	0.115**		0.194***	0.348***		0.310***
	(2.131)		(7.099)	(8.924)		(8.870)
CGOMax52				0.203***	0.427***	0.228***
				(9.138)	(12.41)	(9.284)
CGOMin52				-0.125***	-0.0380	
				(-5.165)	(-1.935)	
Constant	0.00482	0.00280	0.0101**	0.000243	0.0447***	-0.0268***
	(0.791)	(0.497)	(2.014)	(0.0333)	(5.076)	(-3.508)
Observations	5,005	5,005	5,005	5,005	5,005	5,005
R^2	0.673	0.673	0.673	0.668	0.661	0.667
Adj R^2	0.673	0.673	0.673	0.668	0.661	0.667

This table reports results for predictive regressions of CGO on a set of CGO variables. Dependent variable is the percentage difference between the reference point and final price ((Final Price- Reference Price) /Final Price). CGOPurchase is the percentage difference between the reference point and purchase price (Final Price- Purchase Price) /Final Price). CGOMax is the percentage difference between the reference point and maximum price (Final Price- Max) / Final Price). CGOMin is the percentage difference between the reference point and minimum price (Final Price- Min) / Final Price). CGOAverage is the percentage difference between the reference point and average price (Final Price- Average)/ Final Price). CGOMax52 is the percentage difference between the reference point and 52-week maximum price (Final Price- Max52) / Final Price). CGOMin52 is the percentage difference between the reference point and 52-week minimum price (Final Price- Min52) / Final Price). Robust t-statistics in parentheses, clustered by participant
*** p<0.01, ** p<0.05.

Table D: VIF Analysis

Variables	Model A	Model B	Model C	Model D	Model E	Model F
CGO-Purchase	5.90	5.25	4.12	4.51	2.76	4.43
CGO-Max	66.18	3.70	10.61			
CGO-Min	10.42	1.87				
CGO-Average	75.65		13.59	13.91		12.87
CGO-Max52				7.64	2.31	7.20
CGOMin52				1.54	1.42	
Mean VIF	39.54	3.61	9.44	6.90	2.16	8.17

Notes: VIF analysis for models shown in Table C.

Table E- Descriptive Statistics for Independent Variables

	CGO	CGOMax	CGOMin	CGOAverage	CGOMax52	CGOMin52	CGOCom1	CGOCom2
Mean	-0.178	-0.528	0.253	-0.038	-0.310	0.200	-0.167	-0.184
SD	0.754	0.945	0.129	0.265	0.405	0.117	0.518	0.478
Median	-0.013	-0.281	0.240	0.009	-0.194	0.184	-0.044	-0.065
Max	0.665	0.030	0.799	0.564	0.031	0.718	0.322	0.331
Min	-18.841	-25.563	-0.132	-4.048	-7.682	-0.134	-13.139	-10.528
Skew	-8.944	-9.891	0.426	-3.495	-5.423	0.634	-9.107	-7.122

Notes: This table reports summary statistics for CGO variables. All number presented are the time-series average of the cross-sectional statistics. CGO calculated using turnover adjusted purchase price as shown in Equation 5. CGOMax calculated as per CGO but replacing the purchase price with the maximum price. CGOMin calculated as per CGO but replacing the purchase price with the minimum price. CGOAverage calculated as per CGO but replacing the purchase price with the average price. CGOMax52 calculated as per CGO but replacing the purchase price with the 52-week maximum price. CGOMin52 calculated as per CGO but replacing the purchase price with the 52-week minimum price. CGOCom1 calculated as per CGO but replacing the purchase price with Refcom1, calculated as per Equation 6. CGOCom2 calculated as per CGO but replacing the purchase price with Refcom2, calculated as per Equation 7.

Table F- Descriptive Statistics for Control Variables

	Mom	STR	LTR	Avgturn	Mrkcap	BM
Mean	0.164	0.016	0.377	0.003	4.818	-0.693
SD	0.573	0.143	1.014	0.005	1.720	0.976
Median	0.077	0.004	0.187	0.002	4.584	-0.558
Max	10.520	2.379	20.004	0.126	11.603	2.991
Min	-0.866	-0.606	-0.918	0.000	2.108	-7.251
Skew	4.405	3.121	5.673	9.735	0.652	-0.949

Notes: This table reports summary statistics for control variables. All number presented are the time-series average of the cross-sectional statistics. MOM=12month momentum, excluding the last month. STR=Short term reversal, calculated as the return of the last month t. LTR=Long-term reversal, calculated as the return over the last 3 years excluding the last year. AvgTurn is average daily turnover over last year. Mrkcap is the log of market capitalisation in thousands. BM is log of the book to market ratio.

Table G- Regression using CGO Composite Variables (Seasonal Analysis)

VARIABLES	Jan only			Feb-Dec		
	Model A	Model B	Model C	Model D	Model E	Model F
CGO	-0.0285***			0.00811***		
	(-3.629)			(6.336)		
CGOCom1		-0.0390***			0.0138***	
		(-3.388)			(6.578)	
CGOCom2			-0.0386***			0.0138***
			(-3.702)			(6.311)
Mom	0.00124	-0.000107	0.00183	0.00524***	0.00491***	0.00445***
	(0.272)	(-0.0253)	(0.390)	(3.102)	(2.949)	(2.861)
STR	-0.0971***	-0.0997***	-0.0930***	-0.0503***	-0.0506***	-0.0532***
	(-6.394)	(-6.423)	(-6.369)	(-9.177)	(-9.127)	(-9.653)
LTR	-0.00656***	-0.00739***	-0.00815***	-0.000459	-0.000444	-0.000358
	(-3.008)	(-3.356)	(-3.616)	(-0.870)	(-0.872)	(-0.680)
Avgturn	0.349	0.548	0.228	-0.764**	-0.784**	-0.729
	(0.391)	(0.619)	(0.236)	(-2.282)	(-2.183)	(-1.897)
Mrkcap	-0.00918***	-0.00915***	-0.00897***	0.000515	0.000421	0.000348
	(-5.287)	(-5.320)	(-5.155)	(1.217)	(0.999)	(0.835)
BM	0.000126	0.000737	0.000740	0.00222***	0.00213***	0.00208***
	(0.0441)	(0.253)	(0.257)	(2.951)	(2.873)	(2.796)
Constant	0.0754***	0.0740***	0.0725***	0.00899***	0.0101***	0.0106***
	(5.024)	(5.025)	(4.925)	(2.655)	(2.995)	(3.186)
Observations	145,191	145,191	145,191	1,604,775	1,604,775	1,604,775
Number of groups	53	53	53	594	594	594
Average R-Squared	0.116	0.115	0.115	0.0642	0.0644	0.0651

Notes: This table reports results for predictive Fama and MacBeth (1973) regressions of one-month ahead returns on CGO with CGO Composites and a set of control variables. Dependent variable= 1 month return in month t+1. CGO calculated using turnover adjusted purchase price as shown in Equation 5. CGOCom1 calculated as per CGO but replacing the purchase price with Refcom1, calculated as per Equation 6. CGOCom2 calculated as per CGO but replacing the purchase price with Refcom2, calculated as per Equation 7. MOM=12month momentum, excluding the last month. STR=Short term reversal, calculated as the return of the last month t. LTR=Long-term reversal, calculated as the return over the last 3 years excluding the last year. AvgTurn is average daily turnover over last year. Mrkcap is the log of market capitalisation in thousands. BM is log of the book to market ratio. T-statistics in parentheses are Newey-West adjusted. *** p<0.01, ** p<0.05.

Table H- Regression using CGO Composite Variables with Additional Risk Controls

VARIABLES	Model A	Model B	Model C	Model D
CGO	0.00356**	-0.0234***	-0.00733**	
	(2.415)	(-3.109)	(-2.230)	
CGOCom1		0.0401***		0.00247
		(3.778)		(0.385)
CGOCom2			0.0168***	0.00546
			(3.516)	(0.885)
STR	-0.0361***	-0.0353***	-0.0378***	-0.0363***
	(-7.628)	(-7.318)	(-7.804)	(-7.588)
LTR	-0.000392	2.26e-05	-9.48e-05	-0.000338
	(-0.816)	(0.0444)	(-0.190)	(-0.715)
Avgturn	-0.463***	-0.658***	-0.447***	-0.456***
	(-2.988)	(-3.922)	(-2.959)	(-3.125)
Mrkcap	-0.00112***	-0.00118***	-0.00113***	-0.00112***
	(-3.035)	(-3.245)	(-3.156)	(-3.062)
BM	0.00132	0.000862	0.00113	0.0013
	(1.754)	(1.118)	(1.523)	(1.748)
Beta	0.00145	0.00163	0.00139	0.00127
	(1.607)	(1.832)	(1.601)	(1.473)
IVol	-0.155**	-0.116	-0.142	-0.135
	(-2.184)	(-1.609)	(-1.963)	(-1.870)
Disp	-0.00233	-0.00211	-0.00225	-0.00232
	(-1.034)	(-1.001)	(-1.050)	(-1.057)
Constant	0.0219***	0.0231***	0.0223***	0.0217***
	(5.739)	(6.327)	(5.985)	(5.756)
Observations	962,014	962,014	962,014	962,014
R-squared	0.084	0.087	0.088	0.088
Number of groups	491	491	491	491

Notes: This table reports results for predictive Fama and MacBeth (1973) regressions of one-month ahead returns on CGO with CGO Composites and a set of control variables. Dependent variable= 1 month return in month t+1. CGO calculated using turnover adjusted purchase price as shown in Equation 5. CGOCom1 calculated as per CGO but replacing the purchase price with Refcom1, calculated as per Equation 6. CGOCom2 calculated as per CGO but replacing the purchase price with Refcom2, calculated as per Equation 7. MOM=12month momentum, excluding the last month. STR=Short term reversal, calculated as the return of the last month t. LTR=Long-term reversal, calculated as the return over the last 3 years excluding the last year. AvgTurn is average daily turnover over last year. Mrkcap is the log of market capitalisation in thousands. BM is log of the book to market ratio. Beta is calculated from monthly returns over the previous five years [min 2 years]. IVol is idiosyncratic return volatility calculated using the Fama-French 3 factor model over the last year [min 6 months]. Disp is analyst forecast dispersion taken from the I/B/E/S summary file, scaled by the prior year end share price [Jan 1976-Dec 2016]. T-statistics in parentheses are Newey-West adjusted.
*** p<0.01, ** p<0.05.

Table I- Double Cluster Weighted Regression using CGO Composite Variables

VARIABLES	Model A	Model B	Model C	Model D
CGO	0.00157***	-0.00102**	-0.000469	-0.00111**
	(4.474)	(-2.325)	(-1.138)	(-2.491)
CGOCom1		0.00270***		0.00149***
		(6.223)		(3.166)
CGOCom2			0.00229***	0.00141***
			(5.316)	(2.720)
Mom	0.000647**	0.000613	0.000479	0.000525
	(2.006)	(1.892)	(1.437)	(1.579)
STR	-0.00175***	-0.00177***	-0.00195***	-0.00188***
	(-9.512)	(-9.594)	(-10.45)	(-10.01)
LTR	-0.000599***	-0.000512**	-0.000455	-0.000463**
	(-2.582)	(-2.205)	(-1.950)	(-1.980)
Avgturn	0.000185	0.000161	0.000231	0.000200
	(0.635)	(0.553)	(0.803)	(0.698)
Mrkcap	-0.000383	-0.000548	-0.000559	-0.000582
	(-1.220)	(-1.736)	(-1.782)	(-1.849)
BM	0.00116***	0.00108***	0.00108***	0.00107***
	(5.466)	(5.100)	(5.142)	(5.061)
Constant	0.00689	0.00768	0.00799	0.00800
	(1.530)	(1.712)	(1.791)	(1.793)
Observations	1,749,966	1,749,966	1,749,966	1,749,966
R-squared	0.0023	0.0024	0.0024	0.0025

Notes: This table reports results for predictive Double Cluster Weighted regressions of one-month ahead returns on CGO with CGO Composites and a set of control variables. All independent and control variables are decile ranked across the month (1 to 10). Observations within months are weighted to ensure each month receives an equal weight in the regression. Dependent variable= 1 month return in month t+1. CGO calculated using turnover adjusted purchase price as shown in Equation 5. CGOCom1 calculated as per CGO but replacing the purchase price with Refcom1, calculated as per Equation 6. CGOCom2 calculated as per CGO but replacing the purchase price with Refcom2, calculated as per Equation 7. MOM=12month momentum, excluding the last month. STR=Short term reversal, calculated as the return of the last month t. LTR=Long-term reversal, calculated as the return over the last 3 years excluding the last year. AvgTurn is average daily turnover over last year. Mrkcap is the log of market capitalisation in thousands. BM is log of the book to market ratio. T-statistics in parentheses.
*** p<0.01, ** p<0.05.

Table J: Double Sorts by CGOCom1 & CGOCom2

	CGOCom1-1	CGOCom1-2	CGOCom1-3	CGOCom1-4	CGOCom1-5
CGOCom2-1	0.73	0.89	1.21	1.35	1.51
	1.70	3.17**	5.05**	5.83**	6.88**
CGOCom2-2	0.71	0.84	1.11	1.29	1.49
	1.91	3.33**	4.79**	6.28**	7.38**
CGOCom2-3	0.71	0.97	1.19	1.27	1.53
	2.18**	3.91**	5.52**	6.54**	7.42**
CGOCom2-4	0.88	0.88	1.10	1.29	1.72
	2.94**	3.80**	5.38**	6.58**	7.73**
CGOCom2-5	0.96	1.06	1.20	1.19	1.96
	3.47**	4.59**	5.96**	6.32**	7.95**
5-1	**0.23**	**0.17**	**-0.01**	**-0.15**	**0.45**
	0.99	**1.27**	**-0.13**	**-1.35**	**3.36****

This table (Panels A & B) reports returns in double sorted portfolios based on values of CGOCom1 and CGOCom2. At the end of each month, stocks are sorted into 5 portfolios by CGOCom1 and CGOCom2. Stocks in a portfolio are equally weighted. Each portfolio is held for one month and the time series average return in reported in monthly percent. Newey-West corrected t-Statistics are shown below performance.

**=significance at the 5% level.

Table K: Double Sorts by CGOCom2 & CGOCom1

	CGOCom2-1	CGOCom2-2	CGOCom2-3	CGOCom2-4	CGOCom1-5
CGOCom1-1	0.75	0.97	1.10	1.20	1.23
	1.75	3.76**	4.93**	6.08**	6.39**
CGOCom1-2	0.69	0.84	1.08	1.22	1.36
	1.90	3.33**	5.17**	6.32**	6.64**
CGOCom1-3	0.79	0.91	1.12	1.24	1.65
	2.43**	3.81**	5.09**	6.51**	7.81**
CGOCom1-4	0.80	0.91	1.18	1.29	1.74
	2.66**	3.59**	5.56**	6.17**	7.97**
CGOCom1-5	0.89	1.13	1.36	1.49	2.09
	3.05**	4.53**	5.83**	6.75**	8.27**
5-1	**0.14**	**0.16**	**0.26**	**0.30**	**0.87**
	0.64	**1.37**	**2.10****	**2.67****	**6.26****