Internet Appendix for
"Prospect Theory and Stock Market Anomalies"

NICHOLAS BARBERIS, LAWRENCE J. JIN, and BAOLIAN WANG*

This Internet Appendix contains the following sixteen sections:

- Section I: The Probability Weighting Terms
- Section II: The Form of the Objective Function in the Case of Shorting
- Section III: The Generalized Hyperbolic (GH) Skewed t Distribution
- Section IV: Rescaling the Decision Problem
- Section V: Procedure for Computing the Expected Return of an Asset
- Section VI: Stock Market Anomalies
- Section VII: Attitudes to Gambles
- Section VIII: Additional Details about the Equilibrium Structure
- Section IX: Earnings Announcement Analysis
- Section X: Beliefs about Skewness and the Value and Size Anomalies
- Section XI: Investor Heterogeneity
- Section XII: Endogenizing Investors' Initial Holdings
- Section XIII: Testing the Model on An Alternative Set of Anomalies
- Section XIV: Explaining Time-Variation in Anomaly Alphas
- Section XV: Pre- and Post-Publication Alphas
- Section XVI: Interaction Effects

* Citation format: Barberis, Nicholas, Lawrence J. Jin, and Baolian Wang, Internet Appendix for "Prospect Theory and Stock Market Anomalies," *Journal of Finance,* DOI: 10.1111/jofi.13061. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting infor-mation provided by the authors. Any queries (other than missing material) should be directed to the authors of the article.

I. The Probability Weighting Terms

Here, we provide the full expressions for the $dw(P(R_i))$ and $dw(1-P(R_i))$ terms that appear in expressions (11), (16), and (20) in the paper. We can write

$$dw(P(R_i)) = \frac{dw(P(R_i))}{dP(R_i)}\frac{dP(R_i)}{dR_i}dR_i.$$

By differentiating the probability weighting function in equation (6), and using P as shorthand for $P(R_i)$, we can write the right-hand side above as

$$\frac{\delta P^{\delta-1}(P^\delta+(1-P)^\delta)-P^\delta(P^{\delta-1}-(1-P)^{\delta-1})}{(P^\delta+(1-P)^\delta)^{1+\frac{1}{\delta}}}p(R_i)dR_i,$$

where the probability density function $p(R_i)$ is given in equation (12). Similarly,

$$\begin{aligned} dw(1-P(R_i)) &= \frac{dw(1-P(R_i))}{dP(R_i)}\frac{dP(R_i)}{dR_i}dR_i \\ &= -\frac{\delta(1-P)^{\delta-1}(P^\delta+(1-P)^\delta)-(1-P)^\delta((1-P)^{\delta-1}-P^{\delta-1})}{(P^\delta+(1-P)^\delta)^{1+\frac{1}{\delta}}}p(R_i)dR_i. \end{aligned}$$

II. The Form of the Objective Function in the Case of Shorting

In the model, investors are allowed to sell short. If an investor sells short risky asset i, so that $\Theta_i<0$ and hence $\theta_i<0$, the expression in equation (11) is modified as follows:

$$\begin{aligned} & W_0^\alpha \int_{-\infty}^{R_f-\Theta_{i,-1}g_i/\Theta_i}(\Theta_i(R_i-R_f)+\Theta_{i,-1}g_i)^\alpha dw(P(R_i)) \\ & +\lambda W_0^\alpha \int_{R_f-\Theta_{i,-1}g_i/\Theta_i}^{\infty}(\Theta_i(R_f-R_i)-\Theta_{i,-1}g_i)^\alpha dw(1-P(R_i)). \end{aligned}$$

In the same way, the prospect theory terms in equation (20) become

$$\begin{aligned} & \hat{b}\int_{-\infty}^{R_f-\theta_{i,-1}g_i/\theta_i}(\theta_i(R_i-R_f)+\theta_{i,-1}g_i)^\alpha dw(P(R_i)) \\ & +\lambda\hat{b}\int_{R_f-\theta_{i,-1}g_i/\theta_i}^{\infty}(\theta_i(R_f-R_i)-\theta_{i,-1}g_i)^\alpha dw(1-P(R_i)). \end{aligned}$$

III. The Generalized Hyperbolic (GH) Skewed t Distribution

In our model, the vector of asset returns $\widetilde{R} = (\widetilde{R}_1, \ldots, \widetilde{R}_N)'$ follows an N-dimensional GH skewed t distribution. The density function for this distribution is

$$p(R) = \frac{2^{1-\frac{\nu+N}{2}}}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{N}{2}}|S|^{\frac{1}{2}}} \cdot \frac{K_{\frac{\nu+N}{2}}(\sqrt{(\nu + (R-\mu)'S^{-1}(R-\mu))\zeta'S^{-1}\zeta})\exp((R-\mu)'S^{-1}\zeta)}{\left(\sqrt{(\nu + (R-\mu)'S^{-1}(R-\mu))\zeta'S^{-1}\zeta}\right)^{-\frac{\nu+N}{2}}(1+(R-\mu)'S^{-1}(R-\mu)/\nu)^{\frac{\nu+N}{2}}},$$

for $\zeta \neq 0$

$$p(R) = \frac{\Gamma(\frac{\nu+N}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{N}{2}}|S|^{\frac{1}{2}}} . (1+(R-\mu)'S^{-1}(R-\mu)/\nu)^{-\frac{\nu+N}{2}}, \text{ for } \zeta = 0,$$

where $\Gamma(\cdot)$ is the Gamma function and K_l is the modified Bessel function of the second kind with order l.

The above distribution has four parameters: μ, S, ζ, and ν. Here, $\mu = (\mu_1, \ldots, \mu_N)'$, the vector of location parameters, helps to determine the mean of the distribution; S, the dispersion matrix, controls the dispersion in returns; $\zeta = (\zeta_1, \ldots, \zeta_N)'$, the vector of asymmetry parameters, governs the skewness of returns; and ν, a degree of freedom scalar, affects the heaviness of the tails of the distribution. The mean of the distribution is

$$\mu + \frac{\nu}{\nu-2}\zeta.$$

For our computations, we need only the marginal distribution of each asset's return; this is the one-dimensional GH skewed t distribution in equation (12). In that equation, μ_i and ζ_i are the i'th elements of the vectors μ and ζ, respectively, while S_i is the i'th diagonal element of S.

IV. Rescaling the Decision Problem

Substituting the definitions in equation (19) into (16) and multiplying the resulting expression by the exogenous parameter $\Theta_{M,R}^{-1}$, we obtain

$$\theta_i(\mu_i + \frac{\nu\zeta_i}{\nu-2} - R_f) - \frac{\widehat{\gamma}}{2}(\theta_i^2\sigma_i^2 + 2\sum_{j\neq i}\sigma_{ij}\theta_i\theta_{M,j})$$
$$-\lambda\widehat{b}\int_{-\infty}^{R_f-\theta_{i,-1}g_i/\theta_i}(\theta_i(R_f-R_i)-\theta_{i,-1}g_i)^{\alpha}dw(P(R_i))$$

$$-\hat{b}\int_{R_f-\theta_{i,-1}g_i/\theta_i}^{\infty}(\theta_i(R_i-R_f)+\theta_{i,-1}g_i)^{\alpha}dw(1-P(R_i)), \tag{IA1}$$

where

$$\hat{\gamma}=\gamma W_0\Theta_{M,R}, \qquad \hat{b}=bW_0^{\alpha-1}\Theta_{M,R}^{\alpha-1}.$$

It follows that if Θ_i maximizes (16), then $\theta_i=\Theta_i/\Theta_{M,R}$ maximizes (IA1), and conversely that if θ_i maximizes (IA1), then $\Theta_i=\theta_i\Theta_{M,R}$ maximizes (16). Maximizing (16) is therefore equivalent to maximizing (IA1).

The rescaling also allows us to simplify the variance term in the first row of (IA1). Specifically, the quantity

$$\sum_{j\neq i}\theta_{M,j}\sigma_{ij}$$

can be rewritten as

$$-\theta_{M,i}\sigma_i^2+\sum_j\theta_{M,j}\sigma_{ij}=-\theta_{M,i}\sigma_i^2+cov(\widetilde{R}_i,\widetilde{R}_M)=-\theta_{M,i}\sigma_i^2+\beta_i\sigma_M^2,$$

where $\widetilde{R}_M$ is the return on the market portfolio of all risky assets, σ_M is the standard deviation of this return, and β_i is the beta of asset i. Substituting this into (IA1) leads to the expression in (20).

V. Procedure for Computing the Expected Return of an Asset

The equilibrium structure we consider – a bounded-rationality equilibrium with heterogeneous holdings – consists of a location vector $(\mu_1,\ldots,\mu_N)'$ such that, for each asset i, the expression in (20) has either a unique global maximum at $\theta_i^*=\theta_{M,i}$ or else two global maxima at θ_i^* and θ_i^{**} with $\theta_i^*<\theta_{M,i}<\theta_i^{**}$.

We now explain how we compute this heterogeneous-holdings equilibrium. For a given risky asset i, we first check whether investors have identical holdings in the asset, in other words, whether they all hold the same per-capita market supply $\theta_{M,i}$. To do this, we take the first derivative of (20), substitute in $\theta_i=\theta_{M,i}$, and set the resulting expression to zero. This gives[1]

$$0=(\mu_i+\frac{\nu\zeta_i}{\nu-2}-R_f)-\hat{\gamma}\beta_i\sigma_M^2$$

[1]As discussed in Section III of the paper, we take $\theta_{i,-1}=\theta_{M,i}$, which further simplifies (IA2).

$$-\alpha\lambda\widehat{b}\int_{-\infty}^{R_f-\theta_{i,-1}g_i/\theta_{M,i}}(\theta_{M,i}(R_f-R_i)-\theta_{i,-1}g_i)^{\alpha-1}(R_f-R_i)dw(P(R_i))$$
$$-\alpha\widehat{b}\int_{R_f-\theta_{i,-1}g_i/\theta_{M,i}}^{\infty}(\theta_{M,i}(R_i-R_f)+\theta_{i,-1}g_i)^{\alpha-1}(R_i-R_f)dw(1-P(R_i)). \quad \text{(IA2)}$$

We then check whether, for the μ_i that solves (IA2), the function in (20) has a global maximum at $\theta_i = \theta_{M,i}$, as opposed to only a local maximum or a local minimum. If $\theta_i = \theta_{M,i}$ does indeed correspond to a global maximum, then all investors have identical holdings of asset i, each holding the per-capita supply of the asset, namely $\theta_{M,i}$. If the function in (20) does not have a global maximum at this θ_i, then in equilibrium, investors do not have identical holdings of asset i. We must instead look for an equilibrium with heterogeneous holdings.

To find an equilibrium with heterogeneous holdings of asset i, we search for a value of μ_i such that the maximum value of the function in (20) in the range $(-\infty, \theta_{M,i})$, attained at $\theta_i = \theta_i^*$, is equal to the maximum value of the function in the range $(\theta_{M,i}, \infty)$, attained at $\theta_i = \theta_i^{**}$. If we find such a μ_i, then there exists an equilibrium where investors have heterogeneous holdings of the asset, with some investors allocated to $\theta_i = \theta_i^*$ and others allocated to $\theta_i = \theta_i^{**}$. The value of μ_i for which this holds is typically close to the value of μ_i that solves the first-order condition (IA2). Therefore, if we find that, for the value of μ_i that solves (IA2), the function in (20) does not have a global maximum at $\theta_i = \theta_{M,i}$, we search in the neighborhood of that μ_i for an equilibrium with heterogeneous holdings.

VI. Stock Market Anomalies

In this section, we list the predictor variables associated with the 23 anomalies that we study.

IVOL. Idiosyncratic volatility. Standard deviation of the residuals from a firm-level regression of daily stock returns on the daily Fama-French three factors using data from the previous month. See Ang et al. (2006).

SIZE. Market capitalization. The log of the product of price per share and number of shares outstanding, computed at the end of the previous month.

VAL. Book-to-market. The log of book value of equity scaled by market value of equity, computed following Fama and French (1992) and Fama and French (2008); firms with negative book value are excluded from the analysis.

EISKEW. Expected idiosyncratic skewness, computed as in Boyer, Mitton, and Vorkink (2010).

MOM. Momentum. Measured at time t as a stock's cumulative return from the start of month $t-12$ to the end of month $t-2$.

FPROB. Failure probability. Estimated using a dynamic logit model with both accounting and equity market variables as explanatory variables. See Campbell, Hilscher, and Szilagyi (2008).

ZSC. Z-Score. Uses accounting variables to estimate the probability of bankruptcy. See Altman (1968).

NSI. Net stock issuance. Growth rate of split-adjusted shares outstanding over the previous fiscal year. See Stambaugh, Yu, and Yuan (2012).

CEI. Composite equity issuance. Five-year change in number of shares outstanding, excluding changes due to dividends and splits. See Daniel and Titman (2006).

ACC. Accruals. See Sloan (1996).

NOA. Net operating assets. See Hirshleifer et al. (2004).

PROF. Gross profitability. Revenue minus cost of goods sold, divided by assets. See Novy-Marx (2013).

AG. Asset growth. Percentage change in total assets over the previous year. See Cooper, Gulen, and Schill (2008).

ROA. Return on assets. Income before extraordinary items divided by total assets. See Stambaugh, Yu, and Yuan (2012).

INV. Investment to assets. The annual change in gross property, plant, and equipment plus the annual change in inventory, scaled by the lagged book value of assets. See Stambaugh, Yu, and Yuan (2012).

MAX. Measured at time t as a stock's maximum one-day return in month $t-1$. See Bali, Cakici, and Whitelaw (2011).

ORGCAP. Organizational capital. See Eisfeldt and Papanikolaou (2013).

LTREV. Long-term reversal. Measured at time t as a stock's cumulative return from the start of month $t-60$ to the end of month $t-13$.

XFIN. External finance. Total net external financing scaled by total assets. See Bradshaw, Richardson, and Sloan (2006).

STREV. Short-term reversal. Measured at time t as a stock's return in month $t-1$.

DOP. Difference of opinion. The standard deviation of earnings forecasts (unadjusted Institutional Brokers' Estimate System (IBES) file, item STDEV) divided by the absolute value of the consensus forecast (unadjusted file, item MEANEST). We use the forecasts for the current fiscal year. See Diether, Malloy, and Scherbina (2002).

PEAD. Post-earnings announcement drift. Measured as standardized unexpected earnings: the change in quarterly earnings per share from its value four quarters before, divided by the standard deviation of this change in quarterly earnings over the previous eight quarters. See Foster, Olsen, and Shevlin (1984).

CGO. Capital gain overhang. The average percentage capital gain since purchase across investors in a stock; see Grinblatt and Han (2005). These authors compute the gain overhang for stock i as $(P_i - R_i)/P_i$, where P_i is the stock's current price and R_i is investors' average purchase price. We compute it slightly differently, as $(P_i - R_i)/R_i$; this is a more precise match for the capital gain variable g_i in our model.

VII. Attitudes to Gambles

When we implement the model, we set investors' scaled portfolio risk aversion $\widehat{\gamma}$ and scaled weight on the prospect theory term $\widehat{b}$ so as to generate reasonable values for two market features: the equity premium and the level of underdiversification in household portfolios. It is useful to check that the values of these parameters, $(\widehat{\gamma}, \widehat{b}) = (0.6, 0.6)$, also generate sensible attitudes to small- and large-scale gambles.

Suppose that, at time 0, an investor in our economy is endowed with a symmetric gamble: a 50:50 bet to gain \$$M$ or lose \$$M$. What is the largest cash "premium" the investor would be willing to pay to avoid this gamble? For simplicity, we assume that the premium payment and gamble outcomes occur at time 1. If the investor takes the gamble, his utility changes by

$$-\frac{\gamma}{2}M^2 - b(\lambda-1)w\left(\frac{1}{2}\right)M^{\alpha}, \tag{IA3}$$

where $\gamma = \widehat{\gamma}/W_0\Theta_{M,R}$ and $b = \widehat{b}/W_0^{\alpha-1}\Theta_{M,R}^{\alpha-1}$. If he instead pays a premium π at time 1, his utility changes by

$$-\pi - b\lambda\pi^{\alpha}. \tag{IA4}$$

The maximum amount he is willing to pay to avoid the gamble is the value of π that equates (IA3) and (IA4).

Using (IA3) and (IA4), we confirm that, for $\widehat{\gamma} = \widehat{b} = 0.6$ and reasonable values of $\Theta_{M,R}$, the model satisfies the restrictions on attitudes to small- and large-scale gambles proposed by Barberis and Huang (2008b).

We can also examine investor attitudes to the gamble $(-\$M, \frac{1}{2}; \$X, \frac{1}{2})$. If an investor takes this gamble, his utility changes by

$$\frac{X-M}{2} - \gamma\frac{(X+M)^2}{8} - bw\left(\frac{1}{2}\right)(\lambda M^{\alpha} - X^{\alpha}).$$

By setting this expression equal to zero, we can compute how high X needs to be, for a given M, for the investor to be indifferent to the gamble. We confirm that, for $\widehat{\gamma} = \widehat{b} = 0.6$ and reasonable values of $\Theta_{M,R}$, the model captures the restrictions on attitudes to small- and large-scale gambles proposed by Barberis, Huang, and Thaler (2006): for a wide range of wealth levels, the investor rejects the gamble $(-\$500, \frac{1}{2}; \$550, \frac{1}{2})$ and accepts the gamble $(-\$10000, \frac{1}{2}; \$20000000, \frac{1}{2})$.

Finally, we confirm, using a similar methodology, that for $\widehat{\gamma} = \widehat{b} = 0.6$, the model captures the basic experimental evidence in Kahneman and Tversky (1979) that motivates prospect theory, including the rejection of $(-\$100, \frac{1}{2}; \$110, \frac{1}{2})$, the preference for $(-\$1000, \frac{1}{2})$ over $-\$500$, and the preference for $(\$5000, 0.001)$ over $\$5$.

VIII. Additional Details about the Equilibrium Structure

In this section, we present additional information about the equilibrium structure illustrated in Figures 3 and 4 of the paper. We explain why, for some stocks – for example, for stocks in momentum decile 10 – there is no equilibrium with identical holdings for these stocks. We also explain why, for other stocks – for example, stocks in momentum decile 1 – there is no equilibrium with heterogeneous holdings for these stocks. Finally, we contrast the heterogeneous holdings in our model with those that arise in the model of Barberis and Huang (2008a).

No equilibrium with identical holdings for momentum decile 10 stocks. The solid line in Figure 4 shows that, for any stock in momentum decile 10 – for example, for stock 901 –

investors have heterogeneous holdings in the stock. To see why there is no equilibrium with identical holdings in this stock, consider Figure IA.2. The solid line in this figure is the same as the solid line in Figure 4: it plots the objective function in (20) for the location parameter $\mu_{901} = 0.5853$. The dashed line in the figure plots the function in (20) for $\mu_{901} = 0.588$. For this higher value of μ_{901}, the function has a unique global maximum at $\theta_{901} = 0.14 > \theta_{M,901} = 7.26 \times 10^{-4}$. Since demand for stock 901 exceeds supply at this value of μ_{901}, it appears that, to clear the market, we need to lower the value of μ_{901}. However, as we do so, the value of θ_{901} at which the function attains its maximum jumps discontinuously: the dash-dot line, which plots the objective function for a slightly lower value of μ_{901}, namely 0.582, shows that the unique global maximum is now at $\theta_{901} = 0 < \theta_{M,901}$. As such, there is no value of μ_{901} such that the function in (20), for $i = 901$, has a unique global maximum at $\theta_{901} = \theta_{M,901}$. Instead, the market clears only by way of the heterogeneous-holdings structure represented by the solid line.

No equilibrium with heterogeneous holdings for momentum decile 1 stocks. For a stock in momentum decile 1 – for example, for stock 1 – the investors in our model have identical holdings in the stock. Why do heterogeneous holdings not arise in this case? This stock has a loss at time 0 and, in part because of this, has a low expected return. One could then imagine a heterogeneous-holdings structure of the following form. Some investors choose a positive allocation in the stock: since they are in the convex region of the prospect theory value function, they are keen to take risk. Other investors choose a negative allocation in the stock to exploit its low expected return.

Figure 3 in the paper shows that the model comes close to delivering such an equilibrium: the objective function has two local maxima, one in the positive domain and one in the negative domain, which are driven by precisely the forces in the previous paragraph. However, the local maximum in the negative domain has a lower utility level, and so there is a unique global maximum. The reason it has a lower utility level is that, by shorting the stock, the investor is exposing himself to high negative skewness, which he finds aversive.

Can a heterogeneous-holdings structure be achieved by lowering the stock's expected return below the level used in Figure 3? Doing so can certainly give the two local maxima the same utility level. However, the market then no longer clears. As we lower the expected return, the allocation corresponding to the local maximum in the positive domain falls below the market supply. Since the allocations corresponding to the two local maxima are both lower than the market supply, the market no longer clears. In equilibrium, then, the objective function has a unique global maximum and all investors have identical holdings in the stock.

Comparison with the heterogeneous holdings in Barberis and Huang (2008a). An equilibrium with two global maxima also arises in the prospect theory model of Barberis and Huang (2008a). However, the economic forces underlying that equilibrium are different from those that drive the two global maxima in our model. Similar to our model, the model of Barberis and Huang (2008a) features all of the elements of prospect theory; however, it differs from our model in that it does not account for investors' prior gains and losses and assumes broad, rather than narrow, framing. Barberis and Huang (2008a) show that, if a positively skewed security is introduced into the economy, it earns a negative expected excess return. The two global maxima correspond to a zero allocation in this skewed asset and a large undiversified position in it. The economic force underlying the zero allocation is the skewed asset's low expected return; the economic force underlying the high allocation is that, by adding a significant position in the skewed asset to his portfolio, an investor can make his portfolio return more positively skewed, something that, due to probability weighting, he finds attractive.

The mechanism behind the equilibrium structure in Barberis and Huang (2008a) does not arise in our model because it hinges on the broad framing assumption; we assume narrow framing. Importantly, the mechanism in our paper does not arise in Barberis and Huang (2008a). The lower optimum in our Figure 4 reflects investors' desire to preserve a prior gain. This does not apply in the model of Barberis and Huang (2008a) because that model does not account for investors' prior gains and losses.

IX. Earnings Announcement Analysis

Our model performs poorly for some anomalies. One possible reason for this is that some anomalies – we refer to them as "belief-based" anomalies – may be driven not by investor risk attitudes but by incorrect beliefs about firms' future prospects. Below, we identify the anomalies that are more likely to be belief-based by computing, for each anomaly, the fraction of the return spread between deciles 1 and 10 that is earned in the five days around firms' earnings announcements – in other words, on days when investors' erroneous beliefs would be corrected by realized earnings; we call an anomaly belief-based if this fraction is large. We conjecture that the prospect theory model will perform poorly for anomalies that are belief-based and well for anomalies that are not.

We find strong support for this conjecture. For five of the seven anomalies for which the model performs poorly – the value, investment, long-term reversal, accrual, and asset growth anomalies – a large fraction of the anomaly return spread comes around earnings announcement dates. Meanwhile, for several of the anomalies for which the model performs well, little or none of the overall spread comes around these dates.

Our analysis is summarized in Table IA.III. Column (1) reports the monthly return spread for each anomaly – the return difference between decile 10 and decile 1 – in the full sample from 1963 to 2015. Column (2) reports the return spread predicted by our model. Column (3) quantifies the performance of the model – how well it explains an anomaly – as the ratio of column (2) to column (1): a positive number indicates that the model is helpful for thinking about an anomaly, while a negative number means that it is not.

Columns (4) to (8) summarize the earnings announcement results. Due to data availability, we conduct this analysis over a shorter sample that starts in the second quarter of 1983. Column (4) reports the equal-weighted return spread of each anomaly in this sample period. Column (5) reports, for each anomaly, the difference between decile 10 and decile 1 in the cumulative abnormal returns of the underlying firms in the five days around their earnings announcements – put simply, the part of the return spread that comes around earnings announcements. To do this calculation, we obtain earnings announcement dates from the quarterly Compustat and IBES. When the announcement dates in the two databases are not the same, we follow DellaVigna and Pollett (2009) and use the earlier of the two dates.

Column (6) reports the ratio of column (5) to column (4), in other words, the fraction of the return spread that comes in the days around earnings announcements; the higher this number, the more likely that the anomaly is belief-based. We note two outliers in this column, for SIZE and EISKEW; the extreme values for these anomalies are due to their low return spreads in the post-1983 sample period. In column (7), we winsorize the entries for these two anomalies, although our results do not depend on this. Finally, in column (8), we construct a dummy variable that takes the value one if the number in column (7) exceeds 0.5, in other words, if more than half of an anomaly's return spread comes around earnings announcement dates. Using our terminology, a value of one indicates an anomaly that is likely to be belief-based.

Our conjecture is that the prospect theory model will perform better for anomalies that are not belief-based and worse for anomalies that are. Put differently, the numbers in column (3) will be negatively correlated with the numbers in column (7). We find strong support for this conjecture. The correlation is -0.571, which differs from zero in a highly significant way.

X. Beliefs about Skewness and the Value and Size Anomalies

In Section IX of the Internet Appendix, we discuss one reason why the model performs poorly for some anomalies: in most cases, these anomalies appear to be driven primarily by incorrect beliefs about firms' future performance rather than by risk attitudes of the kind captured by prospect theory.

There is another possible explanation for the model's poor performance on some anomalies, one that applies within the context of the model. When computing the model's predictions for anomaly alphas – the predictions in Table III and Figures 5 and 6 – we assume that investors have sensible beliefs about the key model inputs: stocks' volatility, skewness, and gain overhang. However, investors may have incorrect beliefs about these quantities, which may explain the model's poor performance on some anomalies.

We examine this idea in the context of the value and size anomalies. Empirically, value stocks have more volatile returns and more skewed returns than growth stocks, and trade at a larger loss than growth stocks. As a consequence, the model fails to explain the value anomaly: the greater return volatility of value stocks leads the model to predict a higher average return on them, but their higher return skewness and more negative gain overhang lead the model to predict a lower average return on them, and the latter effect dominates.

However, real-world investors may think that growth stocks are more highly skewed than value stocks. If the model accounts for such incorrect beliefs, it will predict a higher (lower) average return on value (growth) stocks, bringing its prediction more in line with the empirical facts. We now examine just how incorrect beliefs about skewness would need to be to generate a substantially positive value premium.

The empirical skewness levels of the typical stocks in the 10 value anomaly deciles are given by the vector

$$skew\_value = [1.85; 1.92; 2.05; 2.43; 2.05; 2.2; 2.34; 2.31; 2.97; 2.66];$$

value stocks (decile 10) have higher return skewness than growth stocks (decile 1). We now construct a vector of incorrect beliefs about the return skewness of the typical stocks in the 10 anomaly deciles, namely,

$$skew\_value\_incorrect = skew\_value + k\ [1; \frac{7}{9}; \frac{5}{9}; \frac{3}{9}; \frac{1}{9}; -\frac{1}{9}; -\frac{3}{9}; -\frac{5}{9}; -\frac{7}{9}; -1],$$

where $k > 0$. This construction leads investors to overestimate the skewness of growth stock returns and to underestimate the skewness of value stock returns while leaving their belief about the return skewness of the average stock unaffected.

We search for the lowest value of k that leads the model to generate a substantially positive value premium, in other words, one that, by the criterion in (28) in the paper, would lead us to say that the model can help explain the value premium. We find that this value of k is approximately 1.8. For this k, investors believe that the return skewness of decile 1 growth stocks is 3.65 and that the return skewness of decile 10 value stocks is 0.86. If investors have sufficiently distorted perceptions of stocks' skewness, then the model can generate a substantially positive value premium. We leave it to future research to determine whether real-world investors hold such distorted beliefs about skewness.

We conduct a similar exercise for the size anomaly. Empirically, the returns of small-cap stocks are much more highly skewed than the returns of large-cap stocks, and this is a major reason why the model incorrectly predicts a negative size premium. However, it is possible that investors underestimate the return skewness of small-cap stocks, and that, once we take this into account, the model will predict a positive size premium. Here, we investigate how distorted investors' perception of the relative return skewness of small-cap and large-cap stocks would need to be to generate a positive size premium.

The empirical skewness levels of the typical stocks in the 10 size anomaly deciles are given by

$$skew\_size = [4.27; 1.89; 1.65; 1.51; 1.28; 1.07; 1.17; 0.96; 0.93; 0.69];$$

small-cap stocks (decile 1) have much higher return skewness than large-cap stocks (decile 10). We now construct a vector of incorrect beliefs about the return skewness of the typical stocks in the 10 anomaly deciles, namely,

$$skew\_size\_incorrect = skew\_size + k\ [1; \frac{7}{9}; \frac{5}{9}; \frac{3}{9}; \frac{1}{9}; -\frac{1}{9}; -\frac{3}{9}; -\frac{5}{9}; -\frac{7}{9}; -1],$$

where $k < 0$. This construction leads investors to underestimate the return skewness of small-cap stocks and to overestimate the return skewness of large-cap stocks while leaving their belief about the return skewness of the average stock unaffected.

We search for the highest value of k that leads the model to predict a positive size premium of 1.5% – a size premium that, by the criterion in (28) in the paper, would lead us to say that the model can help explain the size anomaly. We find that this value of k is approximately -1.9. For this k, investors believe that the return skewness of decile 1 small-cap stocks is 2.37 and that the return skewness of decile 10 large-cap stocks is 2.59. If investors have sufficiently distorted perceptions of stocks' skewness – if they perceive small-cap and large-cap stocks to have fairly similar levels of return skewness – then the model can generate a substantially positive size premium. Again, we leave it to future research to determine whether real-world investors hold such distorted beliefs.

XI. Investor Heterogeneity

For simplicity, we assume that, at time 0, all investors have the same prior gain g_i in stock i. What happens if we allow for heterogeneity across investors in their prior gains? While it is not possible to study heterogeneity in a fully general way, the model does allow us to consider simple forms of heterogeneity. The outcome of this exercise is reassuring: the model predictions are quite robust to heterogeneity in investors' prior gains.

Take, for example, momentum decile 10. The gain overhang of the typical stock in this decile is 30.98%. Accordingly, in our main analysis, we assume that, for this stock, all investors have the same prior gain of 30.98%, an assumption that leads the model to predict an expected return of 8.54% for the stock. Suppose instead that, for this stock, half of the investors have a gain of $30.98 - 10 = 20.98\%$ while the other half have a gain of $30.98 + 10 = 40.98\%$. There is then an equilibrium in which the objective function for the 40.98-gain investors has a unique global maximum at $\theta_i^* = 1.19 \times 10^{-4} < \theta_{M,i} = 7.26 \times 10^{-4}$, while the objective function for the 20.98-gain investors has two global maxima at $\theta_i^* = 6.02 \times 10^{-5} < \theta_{M,i}$ and $\theta_i^{**} = 0.109 > \theta_{M,i}$. In equilibrium, the stock has an expected return of 8.42%. As such, heterogeneity affects the expected return, but in a minor way.

For momentum decile 1, where investors have identical holdings of each stock, the impact of heterogeneity is even smaller. The gain overhang of the typical stock in this decile is -45.02%. In our main analysis, we assume that all investors have the same prior gain of -45.02% in the stock, and this leads to a predicted expected return of -6.80%. Suppose that we instead assume that half of the investors have a gain of $-45.02-10 = -55.02\%$ while the other half have a gain of $-45.02 + 10 = -35.02\%$. In this case, there is an equilibrium in which the objective function of the -55.02-gain investors has a unique global maximum at $\theta_i^* = 2.26 \times 10^{-4} > \theta_{M,i} = 1.86 \times 10^{-4}$, while the objective function of the -35.02-gain investors has a unique global maximum at $\theta_i^* = 1.46 \times 10^{-4} < \theta_{M,i}$. The predicted expected return is *the same* as in the equilibrium with identical investors, namely -6.80%.

The fact that heterogeneity in g_i has a relatively small impact on the model predictions immediately implies that heterogeneity in $\Theta_{i,-1}$, investors' time -1 allocation to risky asset i, will also have a small impact: equation (10) in the paper shows that allowing for heterogeneity in Θ_{-1} is mathematically equivalent to allowing for heterogeneity in g. This helps explain why, as we discuss in Section XII of the Internet Appendix, our conclusions are robust to endogenizing investors' initial holdings.

XII. Endogenizing Investors' Initial Holdings

When we compute the model's predicted alphas, we set investors' initial stock holdings at time -1 equal to stocks' market weights: $\theta_{i,-1} = \theta_{M,i}$ for any stock i. Are our conclusions robust to endogenizing the initial holdings? The answer turns out to be "yes."

To illustrate the results, we take an anomaly that the model can help explain, but not by a wide margin – the gross profitability anomaly. For this anomaly, conditions (28) in the paper are satisfied, but only narrowly so. We want to see if, for this anomaly, conditions (28) continue to hold even after we endogenize the initial holdings.

We start with anomaly decile 1. To endogenize the initial holdings of stock i in this decile, we search for a value of the location parameter μ_i that clears the market at time -1 for the objective function in (20) when $g_i = 0$; the last condition captures the fact that, when an investor first buys the stock, he has no prior gain in it. We find that the market-clearing location parameter is $\mu_i = 0.3682$. For this value of μ_i, the objective function in (20) has two global maxima, $\theta_i^* = 0$ and $\theta_i^{**} = 0.0466$. These maxima straddle the market supply of the stock, allowing us to clear the market by allocating some investors to the first maximum and the remaining investors to the second. These, then, are investors' initial holdings.

We now move to time 0. The typical stock in decile 1 of this anomaly has a capital loss at this time. We find that $\mu_i = 0.3682$ no longer clears the market. The reason is that the investors with an initial holding of 0.0466 now have a substantial prior loss; since they are firmly in the convex region of the prospect theory value function, they want to take a large position in the stock – so large that the market no longer clears. Through a manual search, we find that the market now clears for a different location parameter, $\mu_i = 0.3407$. For this value of μ_i, the objective function in (20) has a single global maximum at $\theta_i = 0$ when $\theta_{i,-1} = 0$, and a single global maximum at $\theta_i = 0.0466$ when $\theta_{i,-1} = 0.0466$. The investors are therefore happy to maintain their time -1 positions and the market clears. The predicted alpha is -1.89%.

We now turn to decile 10. To endogenize the initial holdings of a stock i in this decile, we again search for a value of the location parameter μ_i that clears the market at time -1 for the objective function in (20) when $g_i = 0$. We find that the market clears for $\mu_i = 0.5432$. For this value of μ_i, the objective function has two global maxima at $\theta_i^* = 0$ and $\theta_i^{**} = 0.063$, which become investors' initial holdings. We now move to time 0. At this time, the typical stock in decile 10 has a capital gain. We find that the market clears for the same value of μ_i, namely $\mu_i = 0.5432$. For this μ_i, the investors who had initial holdings of $\theta_{i,-1} = 0.063$ now choose $\theta_i = 0$: they have large gains that they want to secure. Meanwhile, the investors with initial holdings of $\theta_{i,-1} = 0$ are still indifferent between $\theta_i = 0$ and $\theta_i = 0.063$. By allocating some of them to each maximum, we can clear the market. The alpha is 0.83%.

The predicted alpha spread for this anomaly with endogenized initial holdings is therefore 2.73%, and so the conditions in (28) remain satisfied and the model again helps to explain this anomaly.

After conducting this exercise for all 23 anomalies, we obtain the same conclusion as for the benchmark model: the model can again explain 14 of the 23 anomalies – the 14 anomalies in the top group in Table III of the paper. Moreover, the model continues to perform well in terms of average absolute pricing error: it has a pricing error of 0.594, which is again better than the pricing error for the CAPM (0.82) or the three-factor model (0.83) and similar to the pricing error for the four-factor model (0.55). These results reflect the fact that the model's mechanisms, and hence its predictions, are not very sensitive to investors' initial holdings.

XIII. Testing the Model on An Alternative Set of Anomalies

In Section IV of the paper, we use the prospect theory model to generate predictions about the 23 anomalies in Table I and conclude that the model is helpful for thinking about a majority of the anomalies. However, we want to be sure that this conclusion is not particular to this set of anomalies.

After completing the analysis in Section IV, we came across another set of 23 anomalies – one constructed by Novy-Marx and Velikov (2016, NMV) to study transaction costs, in other words, for reasons that have nothing to do with prospect theory. To see if our conclusion in Section IV is robust, we use our model to generate predicted alphas for the 230 NMV anomaly deciles. The results reinforce our conclusion in Section IV. By the criterion in (28) in the paper, the model can help explain 13 of the 23 NMV anomalies, a very similar fraction to that reported in Section IV, where we find that the model can help explain 14 of the 23 anomalies in Table I.

The left column of Table IA.IV lists the 23 NMV anomalies. The NMV anomalies and the anomalies in Table I have 12 anomalies in common; we identify these common anomalies by listing their acronyms in the second column. The third and fourth columns in the table report, for each anomaly, the model-predicted alpha spread – the difference in alphas between anomaly decile 10 and anomaly decile 1, $\alpha^m(10) - \alpha^m(1)$ – and the empirical alpha spread, $\alpha^d(10) - \alpha^d(1)$. The 13 anomalies that the model can help explain, per the criterion in (28), are in the top part of the table. The remaining anomalies are in the lower part of the table.

XIV. Explaining Time-Variation in Anomaly Alphas

To predict the expected return of the typical stock in an anomaly decile, our model requires three main inputs: the stock's return volatility, return skewness, and gain overhang. In Sections III and IV of the paper, we estimate these inputs over the full sample from 1963 to 2015 and then compare the model-predicted alphas to the empirical alphas in this sample period. We find that the model is helpful for thinking about 14 of the 23 anomalies.

Can the model also explain variation over time in the alphas of these 14 anomalies? To answer this question, we divide the full sample into four equal subperiods. For each anomaly decile, we use the procedure described in Section III of the paper to estimate the return volatility, return skewness, and gain overhang of the typical stock in this decile for each of the four subperiods – in other words, using only data from the first subperiod, only data from the second subperiod, and so on. For each subperiod, we use these inputs to generate model-predicted alphas for each anomaly decile in that subperiod.

To see if the model can explain some of the time-variation in the empirical alphas, we proceed as follows. For each of the 14 anomalies, we compute its empirical alpha spread – the difference in the alphas of anomaly decile 10 and anomaly decile 1 – in each of the four subperiods, as well as its model-predicted alpha spread in each of the four subperiods. We then run a regression of the empirical spreads on the model-predicted spreads. To focus on the model's ability to explain time-variation in the alphas, we include anomaly fixed effects. The regression coefficient is positive and highly statistically significant, showing that the model can indeed explain time-variation in anomaly alphas.

Table IA.V reports the model-predicted and empirical alpha spreads for the 14 anomalies in each of the four subperiods. Looking at the table, we can see even without a formal analysis that the model is helpful for thinking about the time-variation in the alphas. For example, several of the anomalies have alpha spreads in the third subperiod that are substantially higher in absolute magnitude, a pattern that the model is often able to capture and one that it ascribes to more extreme values of stocks' return volatility, return skewness, and gain overhang in this subperiod.

XV. Pre- and Post-Publication Alphas

There is currently a lot of interest in out-of-sample anomaly performance – in how anomalies perform after the publication of the journal articles that document them. In this section, we compare the out-of-sample performance of the anomalies for which the prospect theory model performs well with the out-of-sample performance of the anomalies for which it performs poorly. One hypothesis is that, if the anomalies for which the prospect theory model performs well can be said, as a result, to have a sounder theoretical foundation, they should perform better out-of-sample than the anomalies for which the model performs poorly.

To investigate this prediction, for each of the 14 anomalies where the prospect theory model performs well, we identify the date on which the article documenting the anomaly was published. We then compute the anomaly's alpha spread – the difference in alpha between its better-performing extreme decile and its worse-performing extreme decile – in both its pre-publication period and its post-publication period. Averaged across the 14 anomalies, the pre- and post-publication alpha spreads are 13.8% and 7.86%, respectively, as shown in the table below. We do the same computation for the seven anomalies for which the model performs poorly; the pre- and post-publication alpha spreads, averaged across the seven anomalies, are 9.28% and 1.76%, respectively.

From a casual inspection of the table, we see that the anomalies for which the prospect theory model does well "hold up better" in the post-publication period. For these anomalies, their post-publication alpha is on average 63% of their pre-publication alpha, while for the anomalies for which the model does poorly, the post-publication alpha is just 26% of the pre-publication alpha. (To be clear, the 0.63 is the anomaly-by-anomaly ratio of post-publication alpha to pre-publication alpha, averaged across anomalies – it is not 7.86% divided by 13.8%, although the difference is small.)

These results are consistent with the hypothesis we started with: if the anomalies for which the prospect theory model performs well can be said to have a stronger theoretical foundation, they should perform better out of sample. The results are also consistent with the anomalies for which prospect theory performs well being harder to exploit. As discussed in Section V.E of the paper, this is plausible in that the main mispricing generated by prospect theory investors is the overpricing of volatile, skewed, small-cap stocks – mispricing that is hard to arbitrage.

	Pre-publication alpha spread	Post-publication alpha spread	Post/pre ratio	Winsorized ratio
14 anomalies where the model does well	13.8%	7.86%	0.63	0.53
7 anomalies where the model does poorly	9.28%	1.76%	0.26	0.24

We have examined the statistical significance of the above results. The difference between the post-/pre-publication ratios, in other words, the difference between 0.63 and 0.26, is not statistically significant. However, if we apply a reasonable winsorization – one where we constrain the post-publication alpha spread to be at most 100% and at least 0% of the pre-publication alpha spread, a restriction that affects four anomalies, namely EISKEW, ACC, SIZE, and STREV – then the post-/pre-publication ratios become 0.53 and 0.24, which *are* statistically different with a t-statistic of 2.16. In addition, the post-publication alpha spreads for the two groups of anomalies – the 7.86% and the 1.76% – are also statistically different with a t-statistic of 2.56.

XVI. Interaction Effects

In this section, we examine whether the prospect theory model can help explain three interaction effects related to the anomalies in Table I of the paper.

Wang, Yan, and Yu (2017) sort stocks into quintiles based on their gain overhang, where quintile 1 contains the lowest overhang stocks and quintile 5 the highest overhang stocks. They then further sort stocks into quintiles based on idiosyncratic volatility, computed as in Ang et al. (2006), where quintile 5 contains the stocks with the highest volatility. Each stock in the cross-section can then be placed into one of 25 categories, which we label using the notation (i, j), where $i, j \in \{1, 2, 3, 4, 5\}$, so that category (i, j) corresponds to the i'th overhang quintile and j'th idiosyncratic volatility quintile. The authors show that, for stocks in the lowest overhang quintile, there is a negative relationship between idiosyncratic volatility and return: the stocks in $(1, 5)$ have a lower average return and alpha than the stocks in $(1, 1)$. However, for stocks in the highest overhang quintile, there is a positive relationship between idiosyncratic volatility and return: the stocks in $(5, 5)$ have a higher average return and alpha than the stocks in $(5, 1)$.

An et al. (2020) conduct a similar double-sort exercise, this time sorting stocks on their gain overhang and on a measure of skewness, namely expected idiosyncratic skewness. They again create 25 categories, which we again label with the notation (i, j), which denotes the i'th overhang quintile and j'th idiosyncratic skewness quintile. The authors find that, for the quintile of stocks with the lowest overhang, there is a negative relationship between idiosyncratic skewness and return: the stocks in $(1, 5)$ have a lower average return and alpha than the stocks in $(1, 1)$. However, for the stocks in the highest overhang quintile, there is a positive relationship between idiosyncratic skewness and return: the stocks in $(5, 5)$ have a higher average return and alpha than those in $(5, 1)$.

Frazzini (2006) also sorts stocks on two dimensions: their gain overhang and the size of the surprise in their most recent earnings announcement. The 25 resulting categories can again be labeled with the notation (i, j), which now denotes the i'th overhang quintile and j'th earnings surprise quintile. The author shows that a strategy that buys stocks with a positive earnings surprise and high overhang and shorts stocks with a negative earnings surprise and low overhang has a significantly higher alpha than a strategy that buys stocks with a positive earnings surprise and low overhang and shorts stocks with a negative earnings surprise and high overhang. In other words, the alpha of $(5, 5)$ stocks minus the alpha of $(1, 1)$ stocks is significantly higher than the alpha of $(1, 5)$ stocks minus the alpha of $(5, 1)$ stocks.[2]

We now examine whether our model can help explain these empirical patterns. Consider the first pattern – the volatility-overhang interaction described by Wang, Yan, and Yu (2017). We use the methodology of Section III of the paper to compute the return volatility, return skewness, and gain overhang of the typical stock in each of the 25 volatility-overhang categories. We then plug these inputs into the model to see what it predicts for the expected returns and alphas of the 25 categories. We proceed in a similar way for the other two interaction patterns.

We find that the model can help explain all three empirical interactions and illustrate the results in Figure IA.3. The top-left graph in the figure corresponds to the volatility-overhang interaction. The horizontal axis marks the idiosyncratic volatility quintiles, 1 through 5. The lower line in the graph plots the alphas predicted by the model for categories $(1, 1)$ to $(1, 5)$, in other words, the alphas as we increase idiosyncratic volatility within the lowest overhang quintile. The upper line plots the predicted alphas for categories $(5, 1)$ to $(5, 5)$, in other

[2]Frazzini (2006) measures "earnings surprise" using cumulative abnormal stock returns around earnings announcement dates. We measure it instead as the difference between realized earnings and expected earnings derived from a seasonal random walk model. We do not expect this difference in methodology to have a material impact on our results.

words, the alphas as we increase idiosyncratic volatility within the highest overhang quintile. As in the data, the lower line is downward-sloping while the upper line is upward-sloping.

The top-right graph in the figure corresponds to the skewness-overhang interaction. The horizontal axis marks the idiosyncratic skewness quintiles, 1 through 5. The lower line in the graph plots the alphas predicted by the model for categories $(1, 1)$ to $(1, 5)$, in other words, the alphas as we increase idiosyncratic skewness within the lowest overhang quintile. The upper line plots the predicted alphas for categories $(5, 1)$ to $(5, 5)$, in other words, the alphas as we increase idiosyncratic skewness within the highest overhang quintile. As in the data, the lower line is downward-sloping while the upper line is upward-sloping.

Finally, the bottom-left graph in the figure corresponds to the earnings surprise-overhang interaction. The horizontal axis marks the earnings surprise quintiles, 1 through 5. The lower line in the graph plots the alphas predicted by the model for categories $(1, 1)$ to $(1, 5)$, in other words, the alphas as we increase the earnings surprise within the lowest overhang quintile. The upper line plots the predicted alphas for categories $(5, 1)$ to $(5, 5)$, in other words, the alphas as we increase earnings surprise within the highest overhang quintile. As in the data, the difference between the alphas for categories $(5, 5)$ and $(1, 1)$ is much higher than the difference between the alphas for categories $(1, 5)$ and $(5, 1)$.

Wang, Yan, and Yu (2017), An et al. (2020), and Frazzini (2006) suggest, using informal arguments, that the risk attitudes captured by prospect theory can help explain the interaction effects they document. Our analysis provides formal support for this claim. However, it also suggests that the mechanisms through which prospect theory explains the interaction effects may differ from those proposed by these papers.

Consider the top-left graph in Figure IA.3. It is tempting to explain the downward slope of the lower line by noting that, for the negative overhang stocks in categories $(1, 1)$ to $(1, 5)$, investors are in the convex region of the prospect theory value function where they are risk-seeking and, as a result, find volatile stocks appealing. However, our analysis suggests that this is not the main driving force. For the typical stocks in categories $(1, 1)$ to $(1, 5)$, the absolute magnitude of their gain overhang is lower than the standard deviation of their returns; for example, the typical stock in category $(1, 5)$ has a gain overhang of -31% but a standard deviation of 57%. As such, the average investor in this stock is still quite close to the kink in his value function and therefore dislikes volatility.

Our analysis suggests instead that the downward slope of the lower line in the top-left graph arises because the stocks in category $(1, 5)$ tend to have more positively skewed returns and to trade at a larger loss than the stocks in category $(1, 1)$, features that lead investors to require a lower average return on the category $(1, 5)$ stocks. The upward slope of the upper line in the graph is driven primarily by the fact that the stocks in category $(5, 5)$ have more volatile returns than the stocks in category $(5, 1)$; since the average investors in these stocks are to the right of the kink in their value functions, they find volatility unappealing.

Similarly, the downward slope of the lower line in the top-right graph stems largely from the fact that the stocks in category $(1, 5)$ have more positively skewed returns and trade at a larger loss than the stocks in category $(1, 1)$. And the upward slope of the upper line in the graph comes from the fact that the stocks in category $(5, 5)$ are much more volatile than the stocks in category $(5, 1)$.

Our model generates the interaction effect documented by Frazzini (2006) simply because it predicts a higher average return for stocks with a positive gain overhang than for stocks with a negative overhang. To the extent that prospect theory explains the CGO anomaly, it can also explain Frazzini's interaction effect.

REFERENCES

Altman, Edward, 1968, Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy, *Journal of Finance* 23, 589-609.

An, Li, Huijun Wang, Jian Wang, and Jianfeng Yu, 2020, Lottery-related anomalies: The role of reference-dependent preferences, *Management Science* 66, 473-501.

Ang, Andrew, Robert Hodrick, Yuhang Xing, Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

Baele, Lieven, Joost Driessen, Sebastian Ebert, Juan Londono, and Oliver Spalt, 2019, Cumulative prospect theory, option returns, and the variance premium, *Review of Financial Studies* 32, 3667-3723.

Bali, Turan, Nusret Cakici, and Robert Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427-446.

Barberis, Nicholas, and Ming Huang, 2001, Mental accounting, loss aversion, and individual stock returns, *Journal of Finance* 56, 1247-1292.

Barberis, Nicholas, and Ming Huang, 2008a, Stocks as lotteries: The implications of probability weighting for security prices, *American Economic Review* 98, 2066-2100.

Barberis, Nicholas, and Ming Huang, 2008b, The loss aversion / narrow framing approach to the equity premium puzzle, in Rajnish Mehra, ed.: *Handbook of the Equity Risk Premium* (Elsevier).

Barberis, Nicholas, Ming Huang, and Tano Santos, 2001, Prospect theory and asset prices, *Quarterly Journal of Economics* 116, 1-53.

Barberis, Nicholas, Ming Huang, and Richard Thaler, 2006, Individual preferences, monetary gambles, and stock market participation: A case for narrow framing, *American Economic Review* 96, 1069-1090.

Barberis, Nicholas, Lawrence J. Jin, and Baolian Wang, 2021, Prospect theory and stock market anomalies, *Journal of Finance,* forthcoming.

Barberis, Nicholas, Abhiroop Mukherjee, and Baolian Wang, 2016, Prospect theory and stock returns: An empirical test, *Review of Financial Studies* 29, 3068-3107.

Barberis, Nicholas, and Wei Xiong, 2009, What drives the disposition effect? An analysis of a long-standing preference-based explanation, *Journal of Finance* 64, 751-784.

Boyer, Brian, Todd Mitton, and Keith Vorkink, 2010, Expected idiosyncratic skewness, *Review of Financial Studies* 23, 169-202.

Bradshaw, Mark, Scott Richardson, and Richard Sloan, 2006, The relation between corporate financing activities, analysts' forecasts, and stock returns, *Journal of Accounting and Economics* 42, 53-58.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899-2939.

Cooper, Michael, Huseyin Gulen, and Michael Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609-1651.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605-1643.

DellaVigna, Stefano, and Joshua Pollett, 2009, Investor inattention and Friday earnings announcements, *Journal of Finance* 64, 709-749.

Diether, Karl, Christopher Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113-2141.

Eisfeldt, Andrea, and Dimitris Papanikolaou, 2013, Organizational capital and the cross-section of expected returns, *Journal of Finance* 68, 1365-1406.

Fama, Eugene, and Kenneth French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, Eugene, and Kenneth French, 2008, Average returns, B/M, and share issues, *Journal of Finance,* 63, 2971-2995.

Foster, George, Chris Olsen, and Terry Shevlin, 1984, Earnings releases, anomalies, and the behavior of security returns, *Accounting Review* 59, 574-603.

Frazzini, Andrea, 2006, The disposition effect and underreaction to news, *Journal of Finance* 61, 2017-2046.

Grinblatt, Mark, and Bin Han, 2005, Prospect theory, mental accounting, and momentum, *Journal of Financial Economics* 78, 311-339.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics* 38, 297-331.

Ingersoll, Jonathan, 2014, Cumulative prospect theory, aggregation, and pricing, *Critical Finance Review* 4, 1-55.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263-291.

Li, Yan, and Liyan Yang, 2013, Prospect theory, the disposition effect, and asset prices, *Journal of Financial Economics* 107, 715-739.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1-28.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104-147.

Sloan, Richard, 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review* 71, 289-315.

Stambaugh, Robert, Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288-302.

Wang, Huijun, Jinghua Yan, and Jianfeng Yu, 2017, Reference-dependent preferences and the risk-return tradeoff, *Journal of Financial Economics* 123, 395-414.

Table IA.I
The Role of Each Element of Prospect Theory

The left-most column lists the 14 anomalies that the prospect theory model can help explain; the acronyms are defined in Table I of the paper. The next three columns report, for each anomaly, the percentage of the alpha spread – the difference in the model-predicted alphas for deciles 1 and 10 – that is generated by the loss aversion, probability weighting, and diminishing sensitivity elements of prospect theory.

Anomaly	Loss aversion	Probability weighting	Diminishing sensitivity
IVOL	–26	58	68
EISKEW	–62	121	41
MOM	–13	34	79
FPROB	–24	52	73
ZSC	–43	85	58
NSI	–23	60	64
CEI	–23	34	88
PROF	–12	50	62
ROA	–12	33	79
MAX	–11	37	74
XFIN	–4	18	85
DOP	–28	45	84
PEAD	3	6	91
CGO	–22	38	84

Table IA.II
Comparison of Prospect Theory Models

The table lists some papers that present prospect theory models of asset prices and investor behavior and compares their assumptions. The horizonal lines in the table group together models with similar assumptions.

Model	Loss aversion	Diminishing sensitivity	Probability weighting	Accounts for prior gains	Broad (B) or narrow (N) framing	Additional assumptions?
Barberis, Huang, and Santos (2001)	Yes	No	No	Yes	B	Yes
Barberis and Huang (2001)	Yes	No	No	Yes	N	Yes
Grinblatt and Han (2005)	Yes	Yes	No	Yes	N	
Barberis and Xiong (2009)	Yes	Yes	No	Yes	N	
Li and Yang (2013)	Yes	Yes	No	Yes	N	
Barberis and Huang (2008a)	Yes	Yes	Yes	No	B	
Ingersoll (2014)	Yes	Yes	Yes	No	B	
Barberis, Mukherjee, and Wang (2016)	Yes	Yes	Yes	No	N	
Baele et al. (2019)	Yes	No	Yes	No	B	
Barberis, Jin, and Wang (2021)	Yes	Yes	Yes	Yes	N	

Table IA.III
Earnings Announcement Analysis

The left-most column lists 23 anomalies; the acronyms are defined in Table I of the paper. Column (3) reports the fraction of an anomaly's empirical return spread that is explained by the prospect theory model. Column (6) reports the fraction of the return spread that comes around earnings announcement dates; these numbers are winsorized in column (7). The remaining columns are defined in Section IX of the Internet Appendix. The table confirms that the prospect theory model performs better for anomalies where less of the return spread comes around earnings announcements: the numbers in column (3) are negatively correlated with those in column (7).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Anomaly	Empirical return spread	Model predicted return spread	Model performance	Empirical return spread 1983 on	Return spread around earnings	Earnings ratio	Earnings ratio winsorized	Earnings ratio dummy
IVOL	–1.211	–0.591	0.488	–0.671	0.423	–0.63	–0.63	0
SIZE	–0.255	0.334	–1.309	–0.007	–0.283	38.139	1	1
VAL	0.516	–0.627	–1.214	1.507	1.155	0.767	0.767	1
EISKEW	–0.333	–0.447	1.341	–0.087	0.579	–6.635	–1	0
MOM	1.836	1.27	0.692	1.022	0.706	0.69	0.69	1
FPROB	–1.346	–1.109	0.824	–0.778	–0.108	0.139	0.139	0
ZSC	0.776	0.389	0.501	0.543	0.656	1.209	1.209	1
NSI	–0.691	–0.117	0.169	–1.05	–0.714	0.68	0.68	1
CEI	–0.549	–0.079	0.144	–0.883	–0.622	0.705	0.705	1
ACC	–0.607	0.302	–0.498	–0.772	–0.71	0.92	0.92	1
NOA	–0.604	–0.044	0.073	–1.305	–0.382	0.293	0.293	0
PROF	0.423	0.293	0.694	0.656	0.878	1.338	1.338	1
AG	–0.59	0.344	–0.583	–1.304	–0.845	0.648	0.648	1
ROA	0.699	0.708	1.013	1.336	0.579	0.433	0.433	0
INV	–0.591	0.211	–0.357	–1.221	–0.808	0.661	0.661	1
MAX	–0.764	–0.402	0.527	–0.865	0.208	–0.241	–0.241	0
ORGCAP	0.456	–0.093	–0.205	0.738	0.667	0.904	0.904	1
LTREV	–0.423	0.729	–1.722	–0.77	–1.019	1.323	1.323	1
XFIN	–0.687	–0.202	0.294	–1.27	–1.027	0.809	0.809	1
STREV	–0.508	0.565	–1.111	–2.07	–0.259	0.125	0.125	0
DOP	–0.411	–0.02	0.049	–0.761	–0.663	0.872	0.872	1
PEAD	0.563	0.25	0.444	1.43	0.372	0.26	0.26	0
CGO	0.74	1.526	2.063	0.425	–0.332	–0.78	–0.78	0

Table IA.IV
Model Predictions for An Alternative Set of Anomalies

The left-most column lists the 23 anomalies studied by Novy-Marx and Velikov (2016). For those anomalies that are also in the set of anomalies in Table I of the paper, the second column lists their acronyms. The third and fourth columns report, for each anomaly, the alpha spread between anomaly decile 10 and decile 1 predicted by the prospect theory model and the empirical annual value-weighted alpha spread in the 1963 to 2015 sample. The 13 anomalies for which the model performs well are in the upper part of the table; those for which the model performs poorly or does not make a strong prediction are in the lower part.

Anomaly	Acronym	Model alpha spread	Empirical alpha spread
Momentum	MOM	14.52	23.26
Failure probability	FPROB	–13.3	–18.82
Return on book equity		9.31	9.91
Idiosyncratic volatility	IVOL	–8.78	–17.99
Value, profitability, and momentum combination		6.44	17.96
Value and momentum combination		6.39	14.69
Three-day earnings announcement return		5.36	16.5
Piotroski F-score		4.68	3.72
Seasonality		3.99	9.71
Gross profitability	PROF	3.65	6.43
Post-earnings announcement drift	PEAD	3.26	7.12
Net stock issuance	NSI	–2.21	–9.23
Industry momentum		1.77	11.6
Industry relative reversal (low volatility)		–0.53	–13.6
Investment	INV	1.84	–7.8
Market capitalization	SIZE	2.34	–1.76
High-frequency combination		–2.8	21.22
Accrual	ACC	2.89	–8.35
Asset growth	AG	2.93	–8.28
Value and profitability combination		–3.68	9.24
Industry relative reversal		6.35	–7.73
Short-term reversal	STREV	6.79	–3.57
Value	VAL	–6.85	5.79

Table IA.V
Analysis of Time-Variation in Anomaly Alphas

The left-most column lists the 14 anomalies for which the prospect theory model can help explain the anomaly alphas in the full sample from 1963 to 2015; the acronyms are defined in Table I of the paper. The next four columns report, for each anomaly, the model-predicted alpha spread between anomaly decile 10 and anomaly decile 1 for each of four equal subperiods. The last four columns report, for each anomaly, the empirical annual value-weighted alpha spread for each of the four subperiods.

Anomaly	Model Period 1	Model Period 2	Model Period 3	Model Period 4	Data Period 1	Data Period 2	Data Period 3	Data Period 4
IVOL	–8.4	–5.73	–13.53	–6.67	–9.53	–21.55	–37.89	–1.29
EISKEW	–4.80	3.05	–10.03	–6.04	2.19	–1.27	–19.13	–7.61
MOM	14.15	14.77	18.74	10.74	26.24	24.51	31.67	10.39
FPROB	–10.62	–16.77	–16.04	–9.64	–12.25	–20.42	–30.91	–10.62
ZSC	5.34	5.61	8.48	2.28	5.1	15.45	18.74	5.24
NSI	–1.36	0.27	–4.56	–3.66	–8.92	–9.41	–8.2	–10.51
CEI	–2.69	–0.45	–5.06	–2.78	–6.93	–12.2	–12.02	–4.89
PROF	2.26	1.61	4.54	4.9	3.87	0.32	12.56	10.67
ROA	10.98	9.93	11.07	5.72	2.43	10.96	20.13	10.03
MAX	–7.06	–5.18	–12.53	–5.34	–8.33	–13.41	–28.55	–4.86
XFIN	–1.92	–1.71	–6.39	–3.18	–6.2	–10.87	–22.21	–5.76
DOP	–2.93	–2.63	–2.17	–0.62	–9.87	–10.95	–1.85	–11.36
PEAD	2.9	2.85	4.77	2.56	11.89	4.79	5.53	6.6
CGO	17.7	16.56	20.6	15.01	3.0	18.3	20.88	5.05

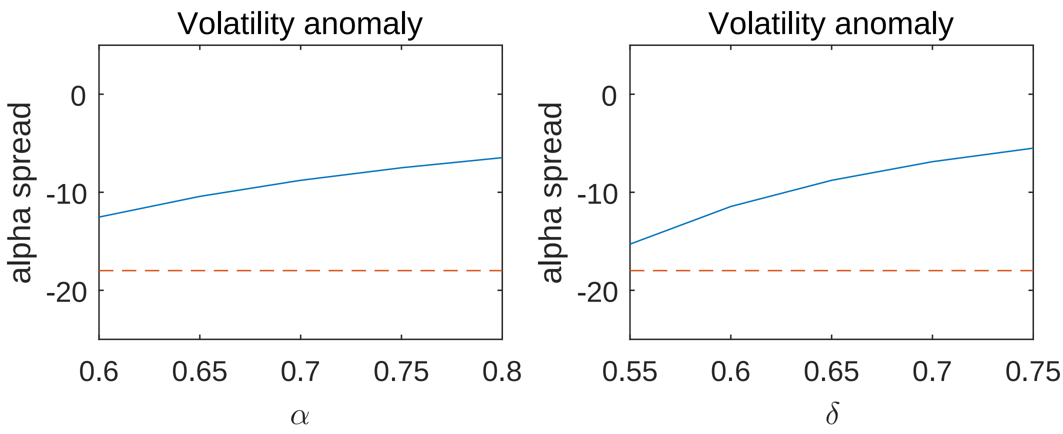

Figure IA.1. Sensitivity of model predictions to preference parameter values. The solid line in the left graph plots the model-predicted alpha spread for the idiosyncratic volatility anomaly for different values of the diminishing sensitivity parameter α; the baseline value is 0.7. The solid line in the right graph plots the predicted alpha spread for this anomaly for different values of the probability-weighting parameter δ; the baseline value is 0.65. The dashed lines in the graphs mark the empirical annual value-weighted alpha spread in the 1963 to 2015 sample period.

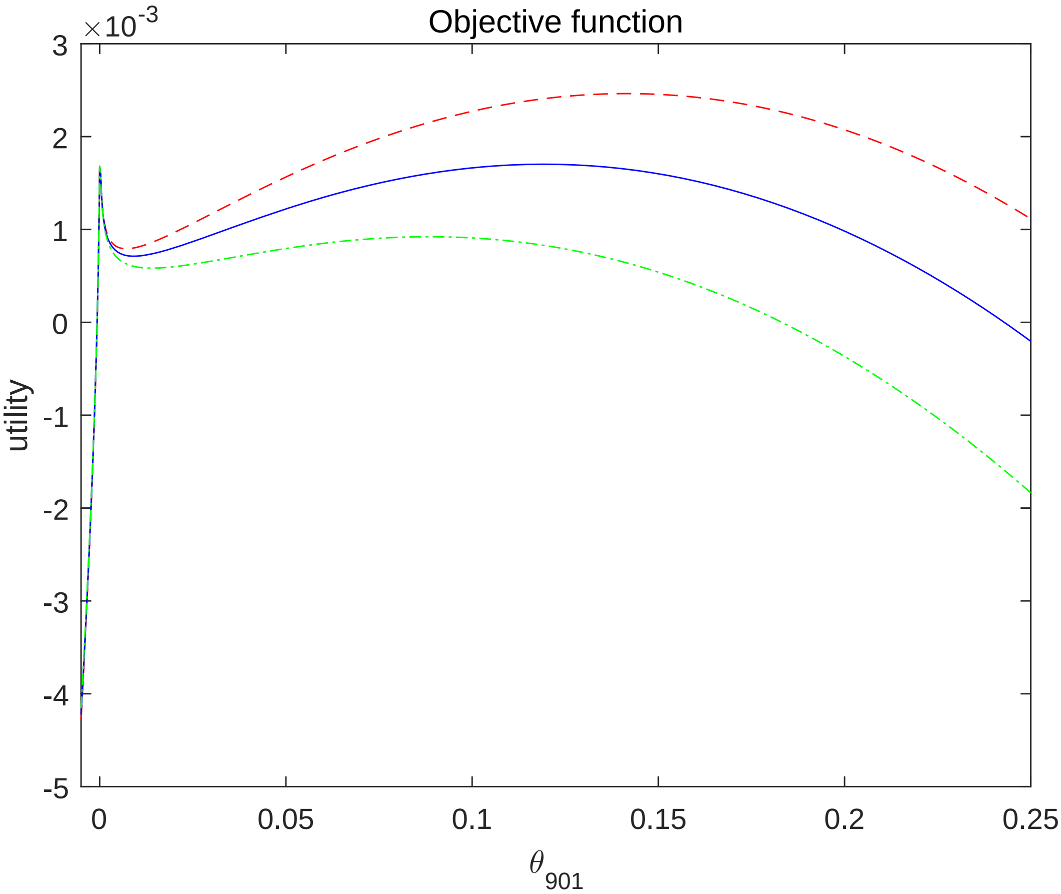

Figure IA.2. Investors' objective function for a stock in momentum decile 10. The solid line plots the value of an investor's objective function in equilibrium as a function of θ_{901}, the (scaled) fraction of the investor's portfolio allocated to stock 901, which belongs to momentum decile 10. The function has two global maxima which straddle the weight of stock 901 in the market portfolio, namely 7.26×10^{-4}. The dashed line plots the objective function for a higher expected return on the stock, while the dash-dot line plots the objective function for a lower expected return. The graph shows that, for any stock in momentum decile 10, investors must have heterogeneous holdings in that stock.

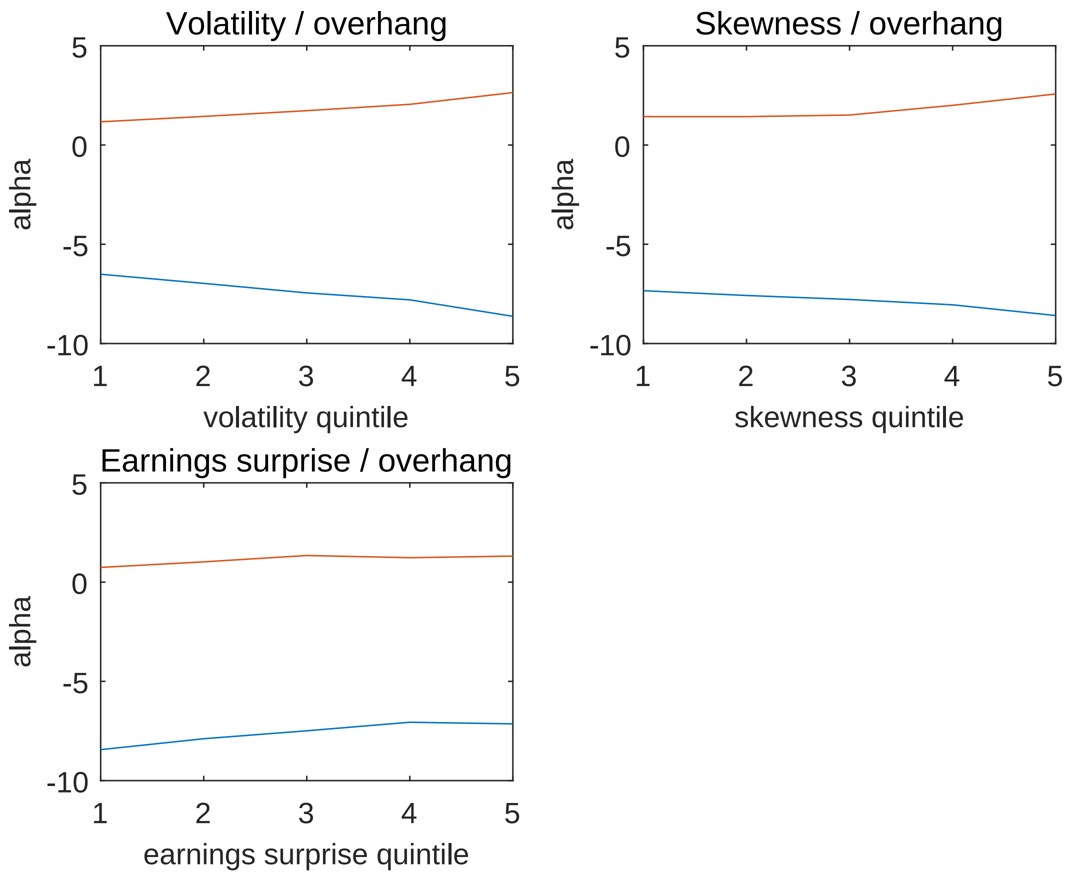

Figure IA.3. Anomaly interaction effects. The graphs show that the prospect theory model can help explain three anomaly interaction effects. The lower lines in the three graphs plot the model-predicted alphas for stocks with large negative overhang and various levels of return volatility, return skewness, and earnings surprise, respectively. The upper lines in the three graphs plot the model-predicted alphas for stocks with large positive overhang and various levels of return volatility, return skewness, and earnings surprise, respectively.